

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2004, 22:01:57 ; Search time 1121 Seconds  
(without alignments)  
11312.106 Million cell updates/sec

Title: US-10-054-841-3  
Perfect score: 2985  
Sequence: 1 gcgcggccgcacgtgtgtgtg.....actgcgcgcgcgcgttgat 2985

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues  
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	ID	Description
1	2985	100.0	2985	3	AAA09801 Human nuc
2	2983.4	99.9	2985	2	AAx32266 Human nuc
3	2973	99.6	2987	3	AAA09802 Human nuc
4	2971.4	99.5	2987	2	AAx32267 Human nuc
5	2901.8	97.2	5216	3	AAA40079 Human nuc
6	2901.8	97.2	5216	9	ADD69756 Human ERR
7	2626	88.0	5221	9	ADD69752 Human ERR
8	2497.8	83.7	3362	9	ADD69754 Human ERR
9	1559	52.2	1615	2	AAV47645 Steroid h
10	1418.2	47.5	1431	3	AAZ95741 Human oes
11	1377	46.1	1377	9	ADC23488 DNA encod
12	1375.4	46.1	1377	3	AAA40078 Human ERR
13	719.6	24.1	1996	3	AAx32265 Human oes
14	715	24.0	2807	2	AAx32265 Human nuc
15	715	24.0	2807	3	AAA09800 Human nuc
16	713.4	23.9	2278	2	AAx99921 Human SBP
17	695	23.3	2153	1	AAx99922 Sequence
18	695	23.3	2153	6	ABZ34975 Human gen
19	695	23.3	2153	6	ABZ35302 Human gen
20	443.2	14.8	536	2	AAV47646 Steroid h
21	427.6	14.3	622	3	AAA09804 Human nuc
22	427.6	14.3	622	2	AAx32269 Nucleotid
23	424.4	14.2	2218	6	ABL64390 Stomach c

24	424.4	14.2	2218	6	ABK83889 Human cDN
25	421.4	14.1	2539	4	AAI58567 Human pol
26	421.4	14.1	2539	8	ADB48544 Novel hum
27	421.4	14.1	4250	6	ABA92902 Human nuc
28	401.6	13.5	2402	6	ABZ34954 Human gen
29	401.6	13.5	2402	7	ABZ76223 Human est
30	399.4	13.4	2249	1	AAx80921 Sequence
31	387.6	13.0	1305	7	ABZ68582 Nucleotid
32	385.8	12.1	2559	4	AAI60353 Human pol
33	360.4	12.1	1599	6	ABV99427 Human NOV
34	356.2	11.9	9193	6	AAD46346 Human nuc
35	248.4	8.3	1131	6	AAx32268 Nucleotid
36	204.6	6.9	403	2	AAx32268 Nucleotid
37	204.6	6.9	403	3	AAA09803 Human nuc
38	178.2	6.0	456	2	AAx99922 Human SBP
39	143.6	4.8	1455	7	AAx99922 Human SBP
40	143.6	4.8	1560	7	AAx99922 Human SBP
41	143.6	4.8	2529	4	AAx99922 Human SBP
42	141.4	4.7	1978	6	ABK89713 Drosophil
43	136.6	4.6	1374	4	ABK89713 Oestrogen
44	136.6	4.6	1380	4	AAx86919 Nucleotid
45	136.6	4.6	1770	9	ADB99351 Fusion pr

ALIGNMENTS

RESULT 1	AAA09801	standard; DNA, 2985 BP.
ID	AAA09801	
XX	AAx32266	
AC	AAx32266	
XX	AAx32266	
DT	05-JUL-2000	(first entry)
XX		
DE	Human nuclear receptor nNR2 gene.	
XX		
KW	Human nuclear receptor protein-2; nNR2; physiological function; ds;	
KW	cell development and differentiation controller; gene expression.	
XX		
OS	Homo sapiens.	
XX		
PN	US6054295-A.	
XX		
PD	25-APR-2000.	
XX		
PF	26-AUG-1998;	98US-00141000.
XX		
PR	27-AUG-1997;	97US-0057090P.
PR	21-OCT-1997;	97US-0062922P.
PR	19-MAR-1998;	98US-0078633P.
XX		
PA	(MERI ) MERCK & CO INC.	
XX		
PI	Chen F;	
XX		
DR	WPI; 2000-328352/28.	
DR	P-PSDB; AAY83823.	
XX		
PT	New polynucleotide encoding human nuclear receptor protein-1 is useful	
PT	for screening compounds that acts as modulators of cell differentiation,	
PT	cell development and physiological function.	
XX		
PS	Disclosure; Fig 5; 58pp; English.	
XX		
CC	This sequence represents the DNA encoding human nuclear receptor protein-	
CC	2 (nNR2). The protein is a cell development and differentiation	
CC	controller which is useful for screening compounds that act as modulators	
CC	of cell differentiation, cell development and physiological functions or	
CC	for treating and diagnosing disorders associated with altered expression	
CC	of the DNA	
XX		
SQ	Sequence 2985 BP; 794 A; 703 C; 605 G; 883 T; 0 U; 0 Other;	

Query Match 100.0%; Score 2985; DB 3; Length 2985;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGGCCGCGCAGTGTGGTGAATTCGGCTTGTCTACTAGGAGAACATTTGTGTTAATTGCA 60  
DB 1 GCGGGCCGCGCAGTGTGGTGAATTCGGCTTGTCTACTAGGAGAACATTTGTGTTAATTGCA 60  
QY 61 CTGTGCTCTGTCAAGAACTTTGATTATAGCTGGGGTGCACAAATATGTTGCCGT 120  
DB 61 CTGTGCTCTGTCAAGAACTTTGATTATAGCTGGGGTGCACAAATATGTTGCCGT 120  
QY 121 CGCACATGATTCGGTAGAATCTTGCTTCTGAATCTTTTCCCTGCACTACGAGGAG 180  
DB 121 CGCACATGATTCGGTAGAATCTTGCTTCTGAATCTTTTCCCTGCACTACGAGGAG 180  
QY 181 AGCTTCTCTGCAAGATGTCAAAAGATCGACACATTGATTCCAGCTGTTCCTTCA 240  
DB 181 AGCTTCTCTGCAAGATGTCAAAAGATCGACACATTGATTCCAGCTGTTCCTTCA 240  
QY 241 TCAAGACGGAACTTCCAGCCCGCTCCCTGACGACAGCGTCAACCAACGAGCCCTG 300  
DB 241 TCAAGACGGAACTTCCAGCCCGCTCCCTGACGACAGCGTCAACCAACGAGCCCTG 300  
QY 301 GTGGCTCTTCAAGCGCCAGTGGAGCTACAGTCAACCATGAATGGCCATCAGAACGAG 360  
DB 301 GTGGCTCTTCAAGCGCCAGTGGAGCTACAGTCAACCATGAATGGCCATCAGAACGAG 360  
QY 361 TTGACTCGGCACTCTCTACCCCTTCTGCTCTATCTCTGGAGGTAGTGGGCTGTCA 420  
DB 361 TTGACTCGGCACTCTCTACCCCTTCTGCTCTATCTCTGGAGGTAGTGGGCTGTCA 420  
QY 421 AACTGTATGATGATGCTCCAGCACCATTTGTTGAAGATCCCGACGACCAAGTGTGA 480  
DB 421 AACTGTATGATGATGCTCCAGCACCATTTGTTGAAGATCCCGACGACCAAGTGTGA 480  
QY 481 TGCTCACTCGATGCCCCAGAGACTGTGTTTAGTGTGTGACATCGCTTCTGGTACC 540  
DB 481 TGCTCACTCGATGCCCCAGAGACTGTGTTTAGTGTGTGACATCGCTTCTGGTACC 540  
QY 541 ACTATGGGGTAGCATCATGTGAAGCCCTGCAAGGCATTTCTCAAGAGACAATTCAAG 600  
DB 541 ACTATGGGGTAGCATCATGTGAAGCCCTGCAAGGCATTTCTCAAGAGACAATTCAAG 600  
QY 601 AATAGAAATACAGCTGCCCTGCCACGAATGAATGAATCACAAAGCGACGTAAT 660  
DB 601 AATAGAAATACAGCTGCCCTGCCACGAATGAATGAATCACAAAGCGACGTAAT 660  
QY 661 CCGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAAGAGGGGTGC 720  
DB 661 CCGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAAGAGGGGTGC 720  
QY 721 GTCTGACAGAGTAGTGAAGTCCGACGAAGTACAAGCGCAGATAGATCGGAGACA 780  
DB 721 GTCTGACAGAGTAGTGAAGTCCGACGAAGTACAAGCGCAGATAGATCGGAGACA 780  
QY 781 GCCCATACCTGAACCCCTCAGCTGTTCAGCCAGCCAAAAGCCATATTAACAAGATTG 840  
DB 781 GCCCATACCTGAACCCCTCAGCTGTTCAGCCAGCCAAAAGCCATATTAACAAGATTG 840  
QY 841 CACATTTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGACCCCTACTGTCC 900  
DB 841 CACATTTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGACCCCTACTGTCC 900  
QY 901 ACAGTGACATCAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGTTA 960  
DB 901 ACAGTGACATCAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGTTA 960  
QY 961 TCATTGGATGGCGAAGCATATTCAGGCTTCTCAAGCTGTCCCTGGCGGACCATGA 1020  
DB 961 TCATTGGATGGCGAAGCATATTCAGGCTTCTCAAGCTGTCCCTGGCGGACCATGA 1020

QY 1021 GCCTTCTGACAGATGCTTGATGGAATTTGATCCTTGGTGTGATATACCGGTCTCTT 1080  
DB 1021 GCCTTCTGACAGATGCTTGATGGAATTTGATCCTTGGTGTGATATACCGGTCTCTT 1080  
QY 1081 CATTGAGATGAATGTCTATATGACAGCAATTAATAATGACGAGACGATCCAAAT 1140  
DB 1081 CATTGAGATGAATGTCTATATGACAGCAATTAATAATGACGAGACGATCCAAAT 1140  
QY 1141 TAGCAGCCTTCTGATCTAATAATATGCTATATCTGACAGCTGGTAAGAAATACAG 1200  
DB 1141 TAGCAGCCTTCTGATCTAATAATATGCTATATCTGACAGCTGGTAAGAAATACAG 1200  
QY 1201 TGAAGCTGAAAAAGAAATTTGTCAACCCCTCAAGCTATAGCTCTTGTATTCAGACT 1260  
DB 1201 TGAAGCTGAAAAAGAAATTTGTCAACCCCTCAAGCTATAGCTCTTGTATTCAGACT 1260  
QY 1261 CCATGACATAGAAGATGTTGAAGCCGTTCAGAAAGCTTCAGAGTGTCTACATGAAGCC 1320  
DB 1261 CCATGACATAGAAGATGTTGAAGCCGTTCAGAAAGCTTCAGAGTGTCTACATGAAGCC 1320  
QY 1321 TGCAGATTTATGAAGTGGCCAGCATGGAAGACCCCTGTCAGCTGGCAAGATGCTGA 1380  
DB 1321 TGCAGATTTATGAAGTGGCCAGCATGGAAGACCCCTGTCAGCTGGCAAGATGCTGA 1380  
QY 1381 TGACACTGCCACTCTGAGGCGACACTCTACCAAGCGGTGACAGATTTCTACAACATCA 1440  
DB 1381 TGACACTGCCACTCTGAGGCGACACTCTACCAAGCGGTGACAGATTTCTACAACATCA 1440  
QY 1441 AACTAGAGGCAAAAGTCCCAATGCACAACTTTTGGAAATGTTGAGGCCAAGTCT 1500  
DB 1441 AACTAGAGGCAAAAGTCCCAATGCACAACTTTTGGAAATGTTGAGGCCAAGTCT 1500  
QY 1501 GACTAAAAAGCTCCCTGGGCTTCCCATCTTCATGTGAAAAAGGAAATAAACCAAG 1560  
DB 1501 GACTAAAAAGCTCCCTGGGCTTCCCATCTTCATGTGAAAAAGGAAATAAACCAAG 1560  
QY 1561 AGTATGTGGAAGAACTTAGAGTTAGTTAAACAACATCAAAAATCAACAGACTGCACTG 1620  
DB 1561 AGTATGTGGAAGAACTTAGAGTTAGTTAAACAACATCAAAAATCAACAGACTGCACTG 1620  
QY 1621 ATAAATTAGCAGCAAGACTATGAAGCAGCTTTCAGATTCCTCATAGGTTCTGATGAGT 1680  
DB 1621 ATAAATTAGCAGCAAGACTATGAAGCAGCTTTCAGATTCCTCATAGGTTCTGATGAGT 1680  
QY 1681 TCTTCTACTTCTCCATCATCTTCTTCTCTTCTCCCAATTTCTCTTCTCTTTA 1740  
DB 1681 TCTTCTACTTCTCCATCATCTTCTTCTCTTCTCTTCTCCCAATTTCTCTTCTCTTTA 1740  
QY 1741 TTTTCTCTCTTCTTCTCTTCTCACTCCCTTATTTCTTCTCTTCTCTTCTCTTCC 1800  
DB 1741 TTTTCTCTCTTCTTCTCTTCTCACTCCCTTATTTCTTCTCTTCTCTTCTCTTCC 1800  
QY 1801 CATTCCTTTATTTCTTCTCCGCTGCTGCTCTCTTCTTCTTCTTCTTCTTCTCTCAT 1860  
DB 1801 CATTCCTTTATTTCTTCTCCGCTGCTGCTCTCTTCTTCTTCTTCTTCTTCTCTCAT 1860  
QY 1861 TCCTCTCTTCTCATCTCTTCCCTTTTCTTAAATTGAATAGCTTTAGTTAAAAAA 1920  
DB 1861 TCCTCTCTTCTCATCTCTTCCCTTTTCTTAAATTGAATAGCTTTAGTTAAAAAA 1920  
QY 1921 AAAAATCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCCCT 1980  
DB 1921 AAAAATCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCCCT 1980  
QY 1981 TTCCTTCTTCTCTTCTTGAACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 2040  
DB 1981 TTCCTTCTTCTCTTCTTGAACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 2040  
QY 2041 CTTTAAAGAGGTCTCTAAGAGAGATGGAAGCCAGCCCTGCAAGAGATGAGA 2100  
DB 2041 CTTTAAAGAGGTCTCTAAGAGAGATGGAAGCCAGCCCTGCAAGAGATGAGA 2100  
QY 2101 TCCATAATATGATGCGAGTGAACCTTATTTGTAACCATACCGTCCCAATGACTAAGGA 2160



Db 2101 TCCATAATATGGATGCCAGTGAACCTTATTTGTGAACCAATACCGTCCCAATGACTAAGGAA 2160  
QY 2161 TCAAAGAGAGAGAACCAAGTCTCTAAAGTACAGTGCACATATACAAATGACTGAGT 2220  
Db 2161 TCAAAGAGAGAGAACCAAGTCTCTAAAGTACAGTGCACATATACAAATGACTGAGT 2220  
QY 2221 GCAGTATGATTTCAATGGGAGCAGCCCTCTAATTAGACAACTTAAGCAACGTTGATCGG 2280  
Db 2221 GCAGTATGATTTCAATGGGAGCAGCCCTCTAATTAGACAACTTAAGCAACGTTGATCGG 2280  
QY 2281 CTGCTTCTTATCATGCTTTTCCATCTAGATCATGATTAAGCCATTGCTTAAATGT 2340  
Db 2281 CTGCTTCTTATCATGCTTTTCCATCTAGATCATGATTAAGCCATTGCTTAAATGT 2340  
QY 2341 TTTTCAAGTCTTCCAGGATTTGTAGTTAGCTACTATGTAATTTTTCAGGGAATAG 2400  
Db 2341 TTTTCAAGTCTTCCAGGATTTGTAGTTAGCTACTATGTAATTTTTCAGGGAATAG 2400  
QY 2401 TTTAAGCTTATTCATTCATGCAATACTAAAGAAATAGAACTGCAATTTTGTCT 2460  
Db 2401 TTTAAGCTTATTCATTCATGCAATACTAAAGAAATAGAACTGCAATTTTGTCT 2460  
QY 2461 GGCCTTGAAACAATTAGCAACAATATGAAGACAATAATGCTGAAGAGATTTTAA 2520  
Db 2461 GGCCTTGAAACAATTAGCAACAATATGAAGACAATAATGCTGAAGAGATTTTAA 2520  
QY 2521 AAATGTTTGTCTCTTCAAAATGAGATTTTGTGACCAAGCTTTACCACTTTTCA 2580  
Db 2521 AAATGTTTGTCTCTTCAAAATGAGATTTTGTGACCAAGCTTTACCACTTTTCA 2580  
QY 2581 CCATTATTAATATGGAATTTAATCTTACTCAAGCAATAGTTGAAGGGAAGTGCATATT 2640  
Db 2581 CCATTATTAATATGGAATTTAATCTTACTCAAGCAATAGTTGAAGGGAAGTGCATATT 2640  
QY 2641 ATCAGGAGCAATTTATGTTGTGTCAGTGTGCTCCCAACATCAATTTCTTAACATG 2700  
Db 2641 ATCAGGAGCAATTTATGTTGTGTCAGTGTGCTCCCAACATCAATTTCTTAACATG 2700  
QY 2701 AGCTCCAGTTTACCTAATGTTCTACTGACACAAAGATGAGATTACCTTACAGTACTC 2760  
Db 2701 AGCTCCAGTTTACCTAATGTTCTACTGACACAAAGATGAGATTACCTTACAGTACTC 2760  
QY 2761 TGAGTAGTCAATATATTAAGCACTGCACATGAGATATAGATCCGTAGAAATGTCAGGAGT 2820  
Db 2761 TGAGTAGTCAATATATTAAGCACTGCACATGAGATATAGATCCGTAGAAATGTCAGGAGT 2820  
QY 2821 GCACCTCTCTACTTGGGAGGTACAAATGCCATATGATTTCTAGCTGCCATGGTGTAGG 2880  
Db 2821 GCACCTCTCTACTTGGGAGGTACAAATGCCATATGATTTCTAGCTGCCATGGTGTAGG 2880  
QY 2881 AATGTGATTAATGCTGTTGGCAAGTCAACAGACCTGCTCAGAAAGAGCTGTGAGCCAG 2940  
Db 2881 AATGTGATTAATGCTGTTGGCAAGTCAACAGACCTGCTCAGAAAGAGCTGTGAGCCAG 2940  
QY 2941 TATTCATTTAAGAGATTCACCACTGCGCGCGCTTGAT 2985  
Db 2941 TATTCATTTAAGAGATTCACCACTGCGCGCGCTTGAT 2985

RESULT 2  
AAAX32266  
ID AAX32266 standard; DNA; 2985 BP.  
XX  
AC AAX32266;  
XX  
DT 15-JUN-1999 (first entry)  
XX  
DE Human nuclear receptor protein nNR2 encoding DNA.  
XX  
KW Nuclear receptor; nNR1; nNR2; cell differentiation; human; ss.  
XX  
OS Homo sapiens.

XX WO9910367-A1.  
PN 04-MAR-1999.  
XX 27-AUG-1998; 98WO-US017826.  
XX 27-AUG-1997; 97US-0057090P.  
PR 21-OCT-1997; 97US-0062902P.  
PR 19-MAR-1998; 98US-0078633P.  
XX (MERI ) MERCK & CO INC.  
PI Chen F;  
XX WPI; 1999-190586/16.  
DR P-PSDB; AAY03837.  
XX  
PT Newly purified DNA polynucleotides encoding human nuclear trans-acting  
PT receptor proteins - useful in the diagnosis, treatment and prophylaxis of  
PT cell differentiation, development and physiological function.  
XX  
PS Claim 31; Page 48-49; 82pp; English.  
XX  
CC The invention relates to DNA molecules encoding human nuclear receptor  
CC (nNR) proteins nNR1 and nNR2. The nNR proteins (including mutants and/or  
CC fragments) form pharmaceutical compositions that are useful in the  
CC diagnosis, treatment and prophylaxis of cell differentiation, development  
CC and physiological function. The proteins are also useful for identifying  
CC downstream target genes and ligands regulating their activity. In  
CC particular, fusion constructs (especially glutathione S-transferase (GST)  
CC -nNR1 and/or GST-nNR2) expressing fusion proteins are useful in screening  
CC for (ant)agonists that are useful as modulators in cell differentiation,  
CC development and physiological function. Antibodies for the nNR proteins  
CC are useful for measuring the levels of these proteins. The present  
CC sequence represents a DNA encoding a human nNR2 protein.  
XX  
SQ Sequence 2985 BP; 794 A; 702 C; 606 G; 883 T; 0 U; 0 Other;

Query Match 99.9%; Score 2983.4; DB 2; Length 2985;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCGGGCCCGCAGTGTGTGGAATTCGGCTTGTCACTAGAGAACATTTGTTAATTGCA 60  
Db 1 GCGGGCCCGCAGTGTGTGGAATTCGGCTTGTCACTAGAGAACATTTGTTAATTGCA 60  
QY 61 CTGTGCTCTGTCAAGGAACCTTGATTTATAGCTGGGGTGCAAAATATGTTGCCGT 120  
Db 61 CTGTGCTCTGTCAAGGAACCTTGATTTATAGCTGGGGTGCAAAATATGTTGCCGT 120  
QY 121 CGCAGATGATTCGTAGAACTTGCCTTCTCTGAATCTTTTCCCTGCACTACGAGAG 180  
Db 121 CGCAGATGATTCGTAGAACTTGCCTTCTCTGAATCTTTTCCCTGCACTACGAGAG 180  
QY 181 AGCTTCTCTGAGATGTCAAAACAAGATGCACATGATTCAGCTGTCTCTTCA 240  
Db 181 AGCTTCTCTGAGATGTCAAAACAAGATGCACATGATTCAGCTGTCTCTTCA 240  
QY 241 TCAAGACGGAACCTTCCAGCCAGCCTCCCTGACGAGACAGCTCAACCAACGACCTG 300  
Db 241 TCAAGACGGAACCTTCCAGCCAGCCTCCCTGACGAGACAGCTCAACCAACGACCTG 300  
QY 301 GTGGCTCTTCAAGCGCAGTGGGAGCTACAGTTCAACCATGATGGCCATCAGAACGAGC 360  
Db 301 GTGGCTCTTCAAGCGCAGTGGGAGCTACAGTTCAACCATGATGGCCATCAGAACGAGC 360  
QY 361 TTGACTGCGCACTCTCTACCCCTTCTGCTCTATCTGGAAGGTAGTGGCTGTCAAGA 420  
Db 361 TTGACTGCGCACTCTCTACCCCTTCTGCTCTATCTGGAAGGTAGTGGCTGTCAAGA 420  
QY 421 AACTGTATGATGACTGCTCCAGCAACATTTGGAAGATCCCGACCAAGTGTGAATACA 480  
Db 421 AACTGTATGATGACTGCTCCAGCAACATTTGGAAGATCCCGACCAAGTGTGAATACA 480

Db 421 AACTGTATGATGACTGCTCCAGACACCAATTGTTGAAGATCCCCAGACCAAGTGTGAATACA 480  
QY 481 TGCTCAACTCGATGCCCAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACC 540  
Db 481 TGCTCAACTCGATGCCCAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACC 540  
QY 541 ACTATGGGGTAGCATCATGTGAAGCCCTGCAAGGCATTCTTCAAGAGGACAAATTCAAGGCA 600  
Db 541 ACTATGGGGTAGCATCATGTGAAGCCCTGCAAGGCATTCTTCAAGAGGACAAATTCAAGGCA 600  
QY 601 ATATAGATACAGCTGCCCTGCCACGAATGTAATGTGAATTCACAAAGCGCAGACGTAAAT 660  
Db 601 ATATAGATACAGCTGCCCTGCCACGAATGTAATGTGAATTCACAAAGCGCAGACGTAAAT 660  
QY 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTTTAAAGTGGGCATGCTGAAGAAGGGGTGC 720  
Db 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTTTAAAGTGGGCATGCTGAAGAAGGGGTGC 720  
QY 721 GTCTTGACAGAGTACGTGAGGTGGGCAAGTACAGCGCAGATAGATGCGGAGACA 780  
Db 721 GTCTTGACAGAGTACGTGAGGTGGGCAAGTACAGCGCAGATAGATGCGGAGACA 780  
QY 781 GCCCATACCTGAACCCCTCAGCTGGTTCAGCCAGCCAAAAAGCCATATTAACAAGATTGCT 840  
Db 781 GCCCATACCTGAACCCCTCAGCTGGTTCAGCCAGCCAAAAAGCCATATTAACAAGATTGCT 840  
QY 841 CACATTTGTTGGTGGCTGAAACCGGAGAAAGATCTATGCCATGCTGACCTTACTGTCCCCG 900  
Db 841 CACATTTGTTGGTGGCTGAAACCGGAGAAAGATCTATGCCATGCTGACCTTACTGTCCCCG 900  
QY 901 ACAGTGACATCAAAAGCCCTCACTACATGTGTGACTTGGCCGACCGAGAGTGTGTTA 960  
Db 901 ACAGTGACATCAAAAGCCCTCACTACATGTGTGACTTGGCCGACCGAGAGTGTGTTA 960  
QY 961 TCATTGGATGGCGAAGCATATTTCCAGGCTTCTTCCAAGCTGTCCCTGGCGGACCAATGA 1020  
Db 961 TCATTGGATGGCGAAGCATATTTCCAGGCTTCTTCCAAGCTGTCCCTGGCGGACCAATGA 1020  
QY 1021 GCCTTCTGCAGAGTGTGTGATGGAATTTGATCCTTGGTGTGCTGATACCGGTCTCTTT 1080  
Db 1021 GCCTTCTGCAGAGTGTGTGATGGAATTTGATCCTTGGTGTGCTGATACCGGTCTCTTT 1080  
QY 1081 CATTTGAGGATGAACCTTGTCTATGCAGCATATATATATGAAGAGACCAAGTCCAAAT 1140  
Db 1081 CATTTGAGGATGAACCTTGTCTATGCAGCATATATATATGAAGAGACCAAGTCCAAAT 1140  
QY 1141 TAGCAGGCTTCTTGTATCTTAAATATATGCTATCCTGCAGCTGGTAAAGAAATACAAGACA 1200  
Db 1141 TAGCAGGCTTCTTGTATCTTAAATATATGCTATCCTGCAGCTGGTAAAGAAATACAAGACA 1200  
QY 1201 TGAAGCTGGAAAAAGAGAATTTGTCAACCTCAAAGCTATAGCTTGTGCTAATTCAAGACT 1260  
Db 1201 TGAAGCTGGAAAAAGAGAATTTGTCAACCTCAAAGCTATAGCTTGTGCTAATTCAAGACT 1260  
QY 1261 CCATGCACATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAGAGATGTCTTACATGAAGCGC 1320  
Db 1261 CCATGCACATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAGAGATGTCTTACATGAAGCGC 1320  
QY 1321 TGCAGGATTAAGAAGCTGGCCAGACATAGGAAGACCCCTGCTGAGCTGGCAAGATGCTGA 1380  
Db 1321 TGCAGGATTAAGAAGCTGGCCAGACATAGGAAGACCCCTGCTGAGCTGGCAAGATGCTGA 1380  
QY 1381 TGACACTGCCACTCTCTGAGGACAGACCTCTACCAAGCCGTGACGATTTCTACAAACATCA 1440  
Db 1381 TGACACTGCCACTCTCTGAGGACAGACCTCTACCAAGCCGTGACGATTTCTACAAACATCA 1440  
QY 1441 AACTAGAAAGCAAAAGTCCCAATGCACAACCTTTTTTGGAAATGTTGAGGGCCAAAGTCT 1500  
Db 1441 AACTAGAAAGCAAAAGTCCCAATGCACAACCTTTTTTGGAAATGTTGAGGGCCAAAGTCT 1500  
QY 1501 GACTAAAAAGCTCCCTGGGCTTCCCATCTCTTCATGTTGAAAAAGGAAAAATAAACCCAG 1560  
Db 1501 GACTAAAAAGCTCCCTGGGCTTCCCATCTCTTCATGTTGAAAAAGGAAAAATAAACCCAG 1560

QY 1561 AGTATGTGAGAGAAACCTTAGAGTTTAGTTAACAAACATCAAAAATCAACAGACTGCAC TG 1620  
Db 1561 AGTATGTGAGAGAAACCTTAGAGTTTAGTTAACAAACATCAAAAATCAACAGACTGCAC TG 1620  
QY 1621 ATATTTAGCAGAGAACTATGAAGCAGCTTTCAGATTCCTCCATAGGTTCTGATGAGT 1680  
Db 1621 ATATTTAGCAGAGAACTATGAAGCAGCTTTCAGATTCCTCCATAGGTTCTGATGAGT 1680  
QY 1681 TCTTTCTACTTTCTCCATCATCTTCTTCCCTTTCTTCCACATTTCTCTTCTTTA 1740  
Db 1681 TCTTTCTACTTTCTCCATCATCTTCTTCCCTTTCTTCCACATTTCTCTTCTTTA 1740  
QY 1741 TTTTTCCTCTTTCTCTTCTTCAACCTCCCTTATTTCTTGTCTTCTTCAATTCCTAGTTC 1800  
Db 1741 TTTTTCCTCTTTCTCTTCTTCAACCTCCCTTATTTCTTGTCTTCTTCAATTCCTAGTTC 1800  
QY 1801 CATTCCTCTTATTTCTTCCCGTCTGCTGCTGCTCTTCTTCTTCTTCTTCTTCTCAT 1860  
Db 1801 CATTCCTCTTATTTCTTCCCGTCTGCTGCTGCTCTTCTTCTTCTTCTTCTTCTCAT 1860  
QY 1861 TCCCTCTTTTCTCATCTTCCCTTTTCTTCTTAAATTTGAATAAGCTTTAGTTAAAAAA 1920  
Db 1861 TCCCTCTTTTCTCATCTTCCCTTTTCTTCTTAAATTTGAATAAGCTTTAGTTAAAAAA 1920  
QY 1921 AAAAACTCCCTTCCCTCTTCCCTTCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCCCT 1980  
Db 1921 AAAAACTCCCTTCCCTCTTCCCTCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCCCT 1980  
QY 1981 TTCTTTCTCTTCTCTTCTTGAACCTTCTTCCATCTTCTTCTTCTTCTTCTGCTGTA 2040  
Db 1981 TTCTTTCTCTTCTCTTCTTGAACCTTCTTCCATCTTCTTCTTCTTCTTCTGCTGTA 2040  
QY 2041 CTTTAAAGAGGTCTCTAAGTGAAGAGATGAAGACCAAGCCCTGCGAAGATGAGA 2100  
Db 2041 CTTTAAAGAGGTCTCTAAGTGAAGAGATGAAGACCAAGCCCTGCGAAGATGAGA 2100  
QY 2101 TCCATATATGATGCGCAGTGAACCTTATTTGTGAACCATACCGTCCCAATGACTAAGGA 2160  
Db 2101 TCCATATATGATGCGCAGTGAACCTTATTTGTGAACCATACCGTCCCAATGACTAAGGA 2160  
QY 2161 TCAAGAGAGAGAACCAAGTCTCTTAAAGTACAGTGAACATATACAATGTAGCTAGT 2220  
Db 2161 TCAAGAGAGAGAACCAAGTCTCTTAAAGTACAGTGAACATATACAATGTAGCTAGT 2220  
QY 2221 GCAATTTAGATTTCAATGGAGCAGCCTCTAATTTAGACAACTTAAGCAACGTTCATCGG 2280  
Db 2221 GCAATTTAGATTTCAATGGAGCAGCCTCTAATTTAGACAACTTAAGCAACGTTCATCGG 2280  
QY 2281 CTGCTTCTTATCATTTGCTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTAATTGT 2340  
Db 2281 CTGCTTCTTATCATTTGCTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTAATTGT 2340  
QY 2341 TTTTTCAGTCTTCCAGGATTTGTTAGTTAGTCTACTATAGTAATTTTTCAGGGAATAG 2400  
Db 2341 TTTTTCAGTCTTCCAGGATTTGTTAGTTAGTCTACTATAGTAATTTTTCAGGGAATAG 2400  
QY 2401 TTTAAGCTTTATTCATTCATGCAATTAAGAGAAATGAATGAATTCGAAATTTGTGCT 2460  
Db 2401 TTTAAGCTTTATTCATTCATGCAATTAAGAGAAATGAATGAATTCGAAATTTGTGCT 2460  
QY 2461 GGCTTGAACAATTAAGAACATTAATGAAGCAAAATGAATCCGTAAGAGAGATTTTAA 2520  
Db 2461 GGCTTGAACAATTAAGAACATTAATGAAGCAAAATGAATCCGTAAGAGAGATTTTAA 2520  
QY 2521 AAATGTTTGTCTTCTTCAAAATGAGATTTTGTACCAAGCTTTACACTTTTCAG 2580  
Db 2521 AAATGTTTGTCTTCTTCAAAATGAGATTTTGTACCAAGCTTTACACTTTTCAG 2580  
QY 2581 CCAATTTATATATGGAATTTAATCTTCAAGCAATAGTTGAAGGGAAGGTGCATATT 2640  
Db 2581 CCAATTTATATATGGAATTTAATCTTCAAGCAATAGTTGAAGGGAAGGTGCATATT 2640

QY 2641 ATCAAGATGCAATTATGTTGTTGCCAGTCTGTGCCCAACATCAATTTCTTAACATG 2700  
Db 2641 ATCAAGATGCAATTATGTTGTTGCCAGTCTGTGCCCAACATCAATTTCTTAACATG 2700  
QY 2701 AGCTTCAGTTTACCTAAATGTTCACTGACACAAAGATGAGATTACACCTACAGTACTC 2760  
Db 2701 AGCTTCAGTTTACCTAAATGTTCACTGACACAAAGATGAGATTACACCTACAGTACTC 2760  
QY 2761 TGAAGTACATATATATTAAGCACTGCACATGAGATATAGATCCGTAGAATTGTGAGAGT 2820  
Db 2761 TGAAGTACATATATATTAAGCACTGCACATGAGATATAGATCCGTAGAATTGTGAGAGT 2820  
QY 2821 GCACCTCTCTACTTGGAGGTACAAATGGCCATATGATTTCTAGCTGCCATGGTGTAGG 2880  
Db 2821 GCACCTCTCTACTTGGAGGTACAAATGGCCATATGATTTCTAGCTGCCATGGTGTAGG 2880  
QY 2881 AATGTGATACTGCTGTTGCAAAAGTCAACAGACCTTGCTCAGAAAGAGCTGTGAGCCAG 2940  
Db 2881 AATGTGATACTGCTGTTGCAAAAGTCAACAGACCTTGCTCAGAAAGAGCTGTGAGCCAG 2940  
QY 2941 TATTCATTTAAGAAATTCACCACTGGCGCGCGCTTGAT 2985  
Db 2941 TATTCATTTAAGAAATTCACCACTGGCGCGCGCTTGAT 2985

RESULT 3

AAA09802 standard; DNA; 2987 BP.

AAA09802;

05-JUL-2000 (first entry)

Human nuclear receptor nNR2-1 variant gene.

Human nuclear receptor protein-2; nNR2; physiological function; ss; cell development and differentiation controller; gene expression.

Homo sapiens.

US6054295-A.

25-APR-2000.

26-AUG-1998; 98US-00141000.

27-AUG-1997; 97US-0057090P.

21-OCT-1997; 97US-0062922P.

19-MAR-1998; 98US-0078633P.

(MERI ) MERCK & CO INC.

Chen F; WPI; 2000-328352/28.

P-PSDB; AAY83824.

New polynucleotide encoding human nuclear receptor protein-1 is useful for screening compounds that acts as modulators of cell differentiation, cell development and physiological function.

Disclosure; Fig 7; 58pp; English.

This sequence represents the DNA encoding variant human nuclear receptor protein 2-1 (nNR2-1). The gene differs from the nNR2 gene (AAA09801) by the insertion of 2 nucleotides at position 1352. The insertion results in the shifting of the reading frame and the creation of a termination codon 33 nucleotides downstream of the insertion site. The translated protein has a C-terminal truncation. The nNR proteins are cell development and differentiation controllers which are useful for screening compounds that act as modulators of cell differentiation, cell development and physiological functions or for treating and diagnosing disorders associated with altered expression of the DNA

XX Sequence 2987 BP; 795 A; 703 C; 606 G; 883 T; 0 U; 0 Other;  
SQ

Query Match 99.6%; Score 2973; DB 3; Length 2987;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2985; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 GCGGGCCGCCAGTGTGTGAATTCGGCTGTCACTAGAGAAACATTTGTGTTAATTGCA 60  
Db 1 GCGGGCCGCCAGTGTGTGAATTCGGCTGTCACTAGAGAAACATTTGTGTTAATTGCA 60  
QY 61 CTGTGCTCTGTCAAGGAACTTTGATTTATAGCTGGGTGCACAATAATGTTGCCGT 120  
Db 61 CTGTGCTCTGTCAAGGAACTTTGATTTATAGCTGGGTGCACAATAATGTTGCCGT 120  
QY 121 GCACATGGAATTCGGTAGAACTTTGCCCTTCTGAATCTTTTCCCTGCACACGAGAG 180  
Db 121 GCACATGGAATTCGGTAGAACTTTGCCCTTCTGAATCTTTTCCCTGCACACGAGAG 180  
QY 181 AGCTTCTCTGCAGAAATGTCAAACAAGATCGACAAATTGATTCACAGCTGTCTCTCA 240  
Db 181 AGCTTCTCTGCAGAAATGTCAAACAAGATCGACAAATTGATTCACAGCTGTCTCTCA 240  
QY 241 TCAAGACGGAACCTTCCAGCCCAAGCCTCCCTGACGGAACAGCGTCAACACCAAGCCTG 300  
Db 241 TCAAGACGGAACCTTCCAGCCCAAGCCTCCCTGACGGAACAGCGTCAACACCAAGCCTG 300  
QY 301 GTGGCTCTTCAAGCGCCAGTGGAGCTACAGTTCAACATGAATGGCCATCAGAACGAGC 360  
Db 301 GTGGCTCTTCAAGCGCCAGTGGAGCTACAGTTCAACATGAATGGCCATCAGAACGAGC 360  
QY 361 TTGACTGGCCACCTCTCTACCCCTTCTGCTCTATCTCGGAGGTAGTGGCTGTCAAGA 420  
Db 361 TTGACTGGCCACCTCTCTACCCCTTCTGCTCTATCTCGGAGGTAGTGGCTGTCAAGA 420  
QY 421 AACTGTATGATGACTGCTCCAGCACCATTTGTAAGATCCCCAGACCAAGTGTGAATACA 480  
Db 421 AACTGTATGATGACTGCTCCAGCACCATTTGTAAGATCCCCAGACCAAGTGTGAATACA 480  
QY 481 TGCTCAACTGCATGCCCAAGAGACTGTGTTTAAAGTGTGTGATGACATCGCTTCTGGTACC 540  
Db 481 TGCTCAACTGCATGCCCAAGAGACTGTGTTTAAAGTGTGTGATGACATCGCTTCTGGTACC 540  
QY 541 ACTATGGGGTAGCATCATGTGAAGCCTGCAAGGCATTCTTCAAGAGACAATTCAAGGCA 600  
Db 541 ACTATGGGGTAGCATCATGTGAAGCCTGCAAGGCATTCTTCAAGAGACAATTCAAGGCA 600  
QY 601 ATATAGAAATACAGCTGCCCTGCCAGAAATGTAATGTAATCACAAAAGCGCAGAGTAAT 660  
Db 601 ATATAGAAATACAGCTGCCCTGCCAGAAATGTAATGTAATCACAAAAGCGCAGAGTAAT 660  
QY 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGCGCATGTGAAGAAAGAGGGGTGC 720  
Db 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGCGCATGTGAAGAAAGAGGGGTGC 720  
QY 721 GTCTTGACAGAGTACGTGAGGTGGGAGAAAGTACAAAGCGCAGAGATAGTCGGAGACA 780  
Db 721 GTCTTGACAGAGTACGTGAGGTGGGAGAAAGTACAAAGCGCAGAGATAGTCGGAGACA 780  
QY 781 GCCCATACCTGAACCCCTCAGCTGTGTTCAAGCCAGCCAAAAGCCATATAACAAGATTGCT 840  
Db 781 GCCCATACCTGAACCCCTCAGCTGTGTTCAAGCCAGCCAAAAGCCATATAACAAGATTGCT 840  
QY 841 CACATTTGTTGTTGGCTGAACCGGAGAAAGTATATGCTGATGCTGATGCTGCTGCTGCT 900  
Db 841 CACATTTGTTGTTGGCTGAACCGGAGAAAGTATATGCTGATGCTGATGCTGCTGCTGCT 900  
QY 901 ACAGTGACATCAAAAGCCCTCAGTACACCTGTGTGACTTGGCCGACGAGAGTGTGTGTTA 960  
Db 901 ACAGTGACATCAAAAGCCCTCAGTACACCTGTGTGACTTGGCCGACGAGAGTGTGTGTTA 960  
QY 961 TCATTGATGGGGAAGCATATTCAGGCTTCTCAGCGCTGTCCCTGGCGACCAAGATGA 1020  
Db 961 TCATTGATGGGGAAGCATATTCAGGCTTCTCAGCGCTGTCCCTGGCGACCAAGATGA 1020



```
Db      961 TCATTGGATGGCGAAGCATAATCCAGGCTTCTCCACCGTGTCCCTGGCGGACCAAGATGA 1020
QY      1021 GCCCTTGCAGAGTGCCTTGATGGAATTTTGATCCCTGGTGTGTATACCAGTCTCTTT 1080
Db      1021 GCGTCTGCAGAGTGTCTTGATGGAATTTTGATCCCTGGTGTGTATACCAGTCTCTTT 1080
QY      1081 CATTTGAGATGAACCTGTCTATGACAGCATTTATATATGACGAAGACCAGTCCAAAT 1140
Db      1081 CATTTGAGATGAACCTGTCTATGACAGCATTTATATATGACGAAGACCAGTCCAAAT 1140
QY      1141 TAGCAGGCGCTTCTTGATCTAAATATATGCTATCCCTGACGCTGTAAAGAAATACAGAGCA 1200
Db      1141 TAGCAGGCGCTTCTTGATCTAAATATATGCTATCCCTGACGCTGTAAAGAAATACAGAGCA 1200
QY      1201 TGAAGCTGGAAGAAAGAAATTTGTACCCCTCAAAGCTATAGCTCTTGCTAATTCAGACT 1260
Db      1201 TGAAGCTGGAAGAAAGAAATTTGTACCCCTCAAAGCTATAGCTCTTGCTAATTCAGACT 1260
QY      1261 CCATGCACATAGAAGATGTTGAAGCCGTTGAGAGCTTCAGATGTCTTACATGAAGCGC 1320
Db      1261 CCATGCACATAGAAGATGTTGAAGCCGTTGAGAGCTTCAGATGTCTTACATGAAGCGC 1320
QY      1321 TGCAGGATTTATGAAGCTGGCCAGCACATG- GAAAGACCTCGTCGAGCTGGCAAGATGCT 1378
Db      1321 TGCAGGATTTATGAAGCTGGCCAGCACATG- GAAAGACCTCGTCGAGCTGGCAAGATGCT 1378
QY      1379 GATGACACTGCCACTCCTGAGGAGACCTCTAACCAAGGCCGTGACGATTTCTACAACAT 1438
Db      1379 GATGACACTGCCACTCCTGAGGAGACCTCTAACCAAGGCCGTGACGATTTCTACAACAT 1438
QY      1439 CAAACTAGGAAGGCAAGTCCCAATGCACAAACTTTTGGAAATGTTGAAGGCCAAGGT 1498
Db      1439 CAAACTAGGAAGGCAAGTCCCAATGCACAAACTTTTGGAAATGTTGAAGGCCAAGGT 1498
QY      1441 CAAACTAGGAAGGCAAGTCCCAATGCACAAACTTTTGGAAATGTTGAAGGCCAAGGT 1500
Db      1441 CAAACTAGGAAGGCAAGTCCCAATGCACAAACTTTTGGAAATGTTGAAGGCCAAGGT 1500
QY      1499 CTGACTAAAGCTCCCTGGGCTTCCCATCTTCATGTGTGAAAAAGGAAAAATAAACCCTA 1558
Db      1501 CTGACTAAAGCTCCCTGGGCTTCCCATCTTCATGTGTGAAAAAGGAAAAATAAACCCTA 1560
QY      1559 AGAGTGATGTGGAAGAACTTAGAGTTAGTTAACAACATCAAAAATCAACAGACTGCAC 1618
Db      1561 AGAGTGATGTGGAAGAACTTAGAGTTAGTTAACAACATCAAAAATCAACAGACTGCAC 1620
QY      1619 TGATTAATTTAGCAGCAGACATATGAAGCAGCTTTCAGATTCCTCCATAGGTTCCGTATGA 1678
Db      1621 TGATTAATTTAGCAGCAGACATATGAAGCAGCTTTCAGATTCCTCCATAGGTTCCGTATGA 1680
QY      1679 GTTCTTCTACTTCTCCATCATCTCTTTCCTCTCTTCCACATTTCTCTTCTCTT 1738
Db      1681 GTTCTTCTACTTCTCCATCATCTCTTTCCTCTCTTCCACATTTCTCTTCTCTT 1740
QY      1739 TAATTTTCTCCTTTCTCTTCACTCCCTTAATTCCTTGTGCTCTTTCATTCCCTAGTT 1798
Db      1741 TAATTTTCTCCTTTCTCTTCACTCCCTTAATTCCTTGTGCTCTTTCATTCCCTAGTT 1800
QY      1799 CCGATTCCTCTTATTTCTCCCGTGTGCGCTGCGCTCTTCTTCTTCTTACCTACTCTC 1858
Db      1801 CCGATTCCTCTTATTTCTCCCGTGTGCGCTGCGCTCTTCTTCTTCTTACCTACTCTC 1860
QY      1859 ATTCTCTCTTTTCTCATCTTCCCTTTTCTTAATTTGAATAGCTTTAGTTAAAA 1918
Db      1861 ATTCTCTCTTTTCTCATCTTCCCTTTTCTTAATTTGAATAGCTTTAGTTAAAA 1920
QY      1919 AAAAAAATCTCTCCCTTCCCGCTTCTTCCCTTCTTCTTCTTCCCTTCTCTTCC 1978
Db      1921 AAAAAAATCTCTCCCTTCCCGCTTCTTCCCTTCTTCTTCTTCCCTTCTCTTCC 1980
QY      1979 CTTTCTTCTCTTCTCTCTGACCTTCTTTCATCTTCTTCTTCTTCTCTCTGCTGCTG 2038
Db      1981 CTTTCTTCTCTTCTCTCTGACCTTCTTTCATCTTCTTCTTCTTCTCTCTGCTGCTG 2040
QY      2039 AACTTTTAAAGAGTCTCTAAGTGAAGAGATGGAAGCCAGCCCTGCAAAAGATGGA 2098
Db      2041 AACTTTTAAAGAGTCTCTAAGTGAAGAGATGGAAGCCAGCCCTGCAAAAGATGGA 2100
```

```
QY      2099 GATCCATAATATGATGCCAGTGAAGTATATTGTGAACCAATACCGTCCCAATGACTTAAGG 2158
Db      2101 GATCCATAATATGATGCCAGTGAAGTATATTGTGAACCAATACCGTCCCAATGACTTAAGG 2160
QY      2159 AATCAAGAGAGAGAAACCAACGTTCTTAAAGTACAGTGCACATATACAATTTGACTGA 2218
Db      2161 AATCAAGAGAGAGAAACCAACGTTCTTAAAGTACAGTGCACATATACAATTTGACTGA 2220
QY      2219 GTGCAGTATTAGATTTTCATGGGAGCAGCCCTCTAATTTAGACACACTTAAAGCAACGTTGCATC 2278
Db      2221 GTGCAGTATTAGATTTTCATGGGAGCAGCCCTCTAATTTAGACACACTTAAAGCAACGTTGCATC 2280
QY      2279 GCGTCTCTTATCATTTGCTTTTCCATCTAGTCAATCAGTTACAGCCATTTGATTCCTTAAT 2338
Db      2281 GCGTCTCTTATCATTTGCTTTTCCATCTAGTCAATCAGTTACAGCCATTTGATTCCTTAAT 2340
QY      2339 GTTTTTCAGTCTTCCAGGTATTTGTTAGTTTGTAGTACTATGTAACTTTTCAGGGAAT 2398
Db      2341 GTTTTTCAGTCTTCCAGGTATTTGTTAGTTTGTAGTACTATGTAACTTTTCAGGGAAT 2400
QY      2399 AGTTAAGCTTTATTCATTTCATGCAATTAATAAGAGAAATAGAAATCTGCAATTTGTG 2458
Db      2401 AGTTAAGCTTTATTCATTTCATGCAATTAATAAGAGAAATAGAAATCTGCAATTTGTG 2460
QY      2459 CTGGCTTTGAACAATTAAGCAATTAATGAAGCAAAATGAATCCTGAAGAAATTTT 2518
Db      2461 CTGGCTTTGAACAATTAAGCAATTAATGAAGCAAAATGAATCCTGAAGAAATTTT 2520
QY      2519 AAAAAATGTTTGTCTTCTTACAAATGGAGATTTTGTACCAAGCTTACCACTTTTC 2578
Db      2521 AAAAAATGTTTGTCTTCTTACAAATGGAGATTTTGTACCAAGCTTACCACTTTTC 2580
QY      2579 AGCCATTATTAATATGGAATTTAACTTACTCAAGCAATAGTTGAAGGGAAGGTGCATA 2638
Db      2581 AGCCATTATTAATATGGAATTTAACTTACTCAAGCAATAGTTGAAGGGAAGGTGCATA 2640
QY      2639 TTATCAGGATGCAATTAATGTTGTGTGCCAGTCTGTGCCAAACATCAATTTCTTAAAC 2698
Db      2641 TTATCAGGATGCAATTAATGTTGTGTGCCAGTCTGTGCCAAACATCAATTTCTTAAAC 2700
QY      2699 TGAGCTCCAGTTTACCTAAATGTTCACTGACACAAAGATGAGATTACACCTTACAGTGAC 2758
Db      2701 TGAGCTCCAGTTTACCTAAATGTTCACTGACACAAAGATGAGATTACACCTTACAGTGAC 2760
QY      2759 TCTGAGTAGTCACATATATTAAGCACTGCACATGAGATATAGATCCGTAGAATTGTACAGA 2818
Db      2761 TCTGAGTAGTCACATATATTAAGCACTGCACATGAGATATAGATCCGTAGAATTGTACAGA 2820
QY      2819 GTGCACCTCTCTACTTTGGGAGGTACAATTGCCATATGAATTTCTAGCTGCCATGTGTGTTA 2878
Db      2821 GTGCACCTCTCTACTTTGGGAGGTACAATTGCCATATGAATTTCTAGCTGCCATGTGTGTTA 2880
QY      2879 GGAATGTATACTGCGCTGTTTGCAAAGTCAACAGACCTTGCCTCAGAAAGAGCTGTGAGCC 2938
Db      2881 GGAATGTATACTGCGCTGTTTGCAAAGTCAACAGACCTTGCCTCAGAAAGAGCTGTGAGCC 2940
QY      2939 AGTATTCATTTAAGAGAAATTCACACACTGCGCGCGCGCTTGAT 2985
Db      2941 AGTATTCATTTAAGAGAAATTCACACACTGCGCGCGCGCTTGAT 2987
```

```
RESULT 4
AAK32267
ID AAK32267 standard; DNA; 2987 BP.
XX
AC AAK32267;
XX
DT 15-JUN-1999 (first entry)
XX
DE Human nuclear receptor protein nNR2-1 encoding DNA.
XX
KW Nuclear receptor; nNR1; nNR2; cell differentiation; human; ss.
```

XX Homo sapiens.  
OS  
XX WO9910367-A1.  
PN  
XX 04-MAR-1999.  
PD  
XX 27-AUG-1998; 98WO-US017826.  
PF  
XX 27-AUG-1997; 97US-0057090P.  
PR 21-OCT-1997; 97US-0062902P.  
PR 19-MAR-1998; 98US-0078633P.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Chen F;  
XX  
DR WPI; 1999-190586/16.  
DR P-PSDB; AAY03838.  
XX  
PT Newly purified DNA polynucleotides encoding human nuclear trans-acting  
PT receptor proteins - useful in the diagnosis, treatment and prophylaxis of  
PT cell differentiation, development and physiological function.  
XX  
PS Disclosure; Fig 7A-C; 82pp; English.

CC The invention relates to DNA molecules encoding human nuclear receptor  
CC (nNR) proteins nNR1 and nNR2. The nNR proteins (including mutants and/or  
CC fragments) form pharmaceutical compositions that are useful in the  
CC diagnosis, treatment and prophylaxis of cell differentiation, development  
CC and physiological function. The proteins are also useful for identifying  
CC downstream target genes and ligands regulating their activity. In  
CC particular, fusion constructs (especially glutathione S-transferase (GST)  
CC -nNR1 and/or GST-nNR2) expressing fusion proteins are useful in screening  
CC for (ant)agonists that are useful as modulators in cell differentiation,  
CC development and physiological function. Antibodies for the nNR proteins  
CC are useful for measuring the levels of these proteins. The present  
CC sequence represents a DNA encoding a human nNR2-1 protein  
XX  
SQ Sequence 2987 BP; 795 A; 702 C; 607 G; 883 T; 0 U; 0 Other;

Query Match 99.5%; Score 2971.4; DB 2; Length 2987;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2984; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 GCGGGCCCGCAGTGTGTGAATTCGGCTGTCTAGAGAGAACATTTGTTAATTGCA 60  
Db 1 GCGGGCCCGCAGTGTGTGAATTCGGCTGTCTAGAGAGAACATTTGTTAATTGCA 60  
QY 61 CTGTGCTCTGTCAAGGAACTTTGATTATAGCTGGGGTGACAAATAATGTTGCCGGT 120  
Db 61 CTGTGCTCTGTCAAGGAACTTTGATTATAGCTGGGGTGACAAATAATGTTGCCGGT 120  
QY 121 CGCACATGATTCGGTAGAATTTGCCCTTCTGAATCTTTTCCCTGCACACAGAGGAAG 180  
Db 121 CGCACATGATTCGGTAGAATTTGCCCTTCTGAATCTTTTCCCTGCACACAGAGGAAG 180  
QY 181 AGCTTCTCTGAGATGTCAACAAGATCGACACATTTGATTCAGAGCTGTCTTCA 240  
Db 181 AGCTTCTCTGAGATGTCAACAAGATCGACACATTTGATTCAGAGCTGTCTTCA 240  
QY 241 TCAAGACGGAACCTTCCAGCCAGCCTCCCTGACGGAACGCTCAACCAAGCCCTG 300  
Db 241 TCAAGACGGAACCTTCCAGCCAGCCTCCCTGACGGAACGCTCAACCAAGCCCTG 300  
QY 301 GTGGCTCTTCAAGCGCCAGTGGAGCTACAGTTCAACCATGAATGGCCATCAAGACGAC 360  
Db 301 GTGGCTCTTCAAGCGCCAGTGGAGCTACAGTTCAACCATGAATGGCCATCAAGACGAC 360  
QY 361 TTGACTCGCACTCTCTACCCCTTCTGCTCTATCCTGGAGAGTAGTGCCCTGTACAGA 420  
Db 361 TTGACTCGCACTCTCTACCCCTTCTGCTCTATCCTGGAGAGTAGTGCCCTGTACAGA 420

QY 421 AACTGATGATGACTGCTCCAGACCAATTTGTGAAGATCCCGACCAAGTGTGAATACA 480  
Db 421 AACTGATGATGACTGCTCCAGACCAATTTGTGAAGATCCCGACCAAGTGTGAATACA 480  
QY 481 TGCTCACTCGATGCCCAAGAGACTGTGTTAGTGTGTGACATCGCTTGTGGTACC 540  
Db 481 TGCTCACTCGATGCCCAAGAGACTGTGTTAGTGTGTGACATCGCTTGTGGTACC 540  
QY 541 ACTATGGGGTAGCATCATGTGAAGCCTGCAAGGCATTTCTCAAGAGACATTCAGGCA 600  
Db 541 ACTATGGGGTAGCATCATGTGAAGCCTGCAAGGCATTTCTCAAGAGACATTCAGGCA 600  
QY 601 ATATAGAATACAGCTGCCCTGCCCAAGATGATGTGAATCAAAAGCCGACGCTAAT 660  
Db 601 ATATAGAATACAGCTGCCCTGCCCAAGATGATGTGAATCAAAAGCCGACGCTAAT 660  
QY 661 CCTGCCAGGCTTGCCCGCTTCAATGAAGTTTAAAGTGGGATGCTGAAAGAGGGGTGC 720  
Db 661 CCTGCCAGGCTTGCCCGCTTCAATGAAGTTTAAAGTGGGATGCTGAAAGAGGGGTGC 720  
QY 721 GTCTTGACAGAGTACGTGGAGGTGCGCAGAGTACAAAGCCAGATAGATGCGGAGACA 780  
Db 721 GTCTTGACAGAGTACGTGGAGGTGCGCAGAGTACAAAGCCAGATAGATGCGGAGACA 780  
QY 781 GCCCATACCTGAACCTTCAGCTGCTTACGCCAGCCAAAAGCCATATAACAAGATTGCT 840  
Db 781 GCCCATACCTGAACCTTCAGCTGCTTACGCCAGCCAAAAGCCATATAACAAGATTGCT 840  
QY 841 CACATTTGTTGGTGTGTAACCGGAGAGATCTATGCCATGCTGACCTACTGTCCCG 900  
Db 841 CACATTTGTTGGTGTGTAACCGGAGAGATCTATGCCATGCTGACCTACTGTCCCG 900  
QY 901 ACAGTGACATCAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGTTA 960  
Db 901 ACAGTGACATCAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGTTA 960  
QY 961 TCATTGGATGGCGGAGACATATTCAGAGCTTCTCCAGCTGTCCCTGGCGGACAGATGA 1020  
Db 961 TCATTGGATGGCGGAGACATATTCAGAGCTTCTCCAGCTGTCCCTGGCGGACAGATGA 1020  
QY 1021 GCCTTCTGACAGTGTGTTGATGGAATTTGATTCCTTGTGCTATACCGGTCTTT 1080  
Db 1021 GCCTTCTGACAGTGTGTTGATGGAATTTGATTCCTTGTGCTATACCGGTCTTT 1080  
QY 1081 CATTGAGGAGAACTTGTCTATGACAGACGATATATAATGACGAAGACCAATCCAAT 1140  
Db 1081 CATTGAGGAGAACTTGTCTATGACAGACGATATATAATGACGAAGACCAATCCAAT 1140  
QY 1141 TAGCAGGCTTTTGAATCTAATAATGCTATPCTGACAGCTGTAAAGAAATCAAGACA 1200  
Db 1141 TAGCAGGCTTTTGAATCTAATAATGCTATPCTGACAGCTGTAAAGAAATCAAGACA 1200  
QY 1201 TGAAGCTGAAAAAGAAAGAAATTTGTCAACCTTAAAGCTATAGCTCTTGTAAATTCAGACT 1260  
Db 1201 TGAAGCTGAAAAAGAAAGAAATTTGTCAACCTTAAAGCTATAGCTCTTGTAAATTCAGACT 1260  
QY 1261 CCATGACATAGAAAGATGTGAAGCCGTTCAAGAGCTTCAGAGATGTTAATGAAGCGC 1320  
Db 1261 CCATGACATAGAAAGATGTGAAGCCGTTCAAGAGCTTCAGAGATGTTAATGAAGCGC 1320  
QY 1321 TGCAGGATTATGAAGCTGGCCAGACATG--GAAGACCCTCGTGAAGCTGGCAAGATGCT 1378  
Db 1321 TGCAGGATTATGAAGCTGGCCAGACATG--GAAGACCCTCGTGAAGCTGGCAAGATGCT 1378  
QY 1379 GATGACACTGCGACTCCTGAGGACAGCTTACCAAGGCGGTGACAGATTCTACAACAT 1438  
Db 1379 GATGACACTGCGACTCCTGAGGACAGCTTACCAAGGCGGTGACAGATTCTACAACAT 1438  
QY 1439 CAAACTAGAAAGCAAGTCCCAATGCAAACTTTTGTGAAATGTGGAAGCCCAAGGT 1498  
Db 1439 CAAACTAGAAAGCAAGTCCCAATGCAAACTTTTGTGAAATGTGGAAGCCCAAGGT 1498  
QY 1499 CTGACTAAAGCTCCCTGGGCTTCCCATCTTATGTTGAAAAAGGGAATAAACCCA 1558  
Db 1499 CTGACTAAAGCTCCCTGGGCTTCCCATCTTATGTTGAAAAAGGGAATAAACCCA 1558





CC having lipid metabolism regulatory and fatty acid synthesis functions.  
CC The product of the invention has cytosolic and antiarteriosclerotic  
CC activity. The protein and gene are useful as drugs and in developing  
CC drugs for treatment of e.g. arteriosclerosis and diabetes. The protein is  
CC obtained by cloning human brain-originated cDNA library. This sequence  
CC encodes the human brain ERgamma protein which is described in the method  
CC of the invention

XX Sequence 5216 BP; 1568 A; 1019 C; 974 G; 1655 T; 0 U; 0 Other;

Query Match 97.2%; Score 2901.8; DB 3; Length 5216;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2929; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 17 GTGGAATTGCGCTTGCTGCTAGGAGAACATTTGTGTTAATTGCACTGTGCTCTGTCAAG 76  
Db 46 GTGCTTATAGGCTTGCTGCTAGGAGAACATTTGTGTTAATTGCACTGTGCTCTGTCAAG 105  
QY 77 AAACCTTGATTATAGCTGGGGTGCAACAATAATGTTGCCGGTCCGACATGGATTGGT 136  
Db 106 AAACCTTGATTATAGCTGGGGTGCAACAATAATGTTGCCGGTCCGACATGGATTGGT 165  
QY 137 AGAAGCTTGGCTTCCCTGAATCTTTTCCCTGCACTACGAGAGAGAGCTTCTCTGCAAGT 196  
Db 166 AGAAGCTTGGCTTCCCTGAATCTTTTCCCTGCACTACGAGAGAGAGCTTCTCTGCAAGT 225  
QY 197 GTCAACAAGAATCGACACATTTGATTCAGCTGTTCGCTTCATCAAGACGGAACCTTC 256  
Db 226 GTCAACAAGAATCGACACATTTGATTCAGCTGTTCGCTTCATCAAGACGGAACCTTC 285  
QY 257 CAGCCCAAGCCTCCCTGACGGACAGCGTCAACCAACCAAGCCTGTGGCTCTTCAAGCGC 316  
Db 286 CAGCCCAAGCCTCCCTGACGGACAGCGTCAACCAACCAAGCCTGTGGCTCTTCAAGCGC 345  
QY 317 CAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGAATTGACTGCCACTCT 376  
Db 346 CAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGAATTGACTGCCACTCT 405  
QY 377 CTACCCCTTGCTCTCTATCCTGGAGGTAGTGGCCTGTCAAGAACTGTATGATGACTG 436  
Db 406 CTACCCCTTGCTCTCTATCCTGGAGGTAGTGGCCTGTCAAGAACTGTATGATGACTG 465  
QY 437 CTCACGACCATTTGTAAGATCCCCAGACCAAGTGTGAATACATGCTCACTCGATGCC 496  
Db 466 CTCACGACCATTTGTAAGATCCCCAGACCAAGTGTGAATACATGCTCACTCGATGCC 525  
QY 497 CAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTAACCACTATGGGGTAGCATC 556  
Db 526 CAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTAACCACTATGGGGTAGCATC 585  
QY 557 ATGTGAAGCCTGCAAGGCATTTCTCAAGAGGACAATTCAAGGCAATATAGAAATACGCTG 616  
Db 586 ATGTGAAGCCTGCAAGGCATTTCTCAAGAGGACAATTCAAGGCAATATAGAAATACGCTG 645  
QY 617 CCCTGCCACGAATGAATGTAATCACAAAAGCGCAGAGCTAAATCCTGCCAGGCTTCCG 676  
Db 646 CCCTGCCACGAATGAATGTAATCACAAAAGCGCAGAGCTAAATCCTGCCAGGCTTCCG 705  
QY 677 CTTTCATGAAGTGTTTAAAGTGGGCATGCTGAAAAGAGGGGTGCGCTTTCAGACAGTACG 736  
Db 706 CTTTCATGAAGTGTTTAAAGTGGGCATGCTGAAAAGAGGGGTGCGCTTTCAGACAGTACG 765  
QY 737 TGGAGGTGCGCAGAGTACAGCGCAGAGTAGATGCGGAGAACAGCCCATACCTGAACCC 796  
Db 766 TGGAGGTGCGCAGAGTACAGCGCAGAGTAGATGCGGAGAACAGCCCATACCTGAACCC 825  
QY 797 TCAGCTGTGTCAGCCAGCCAAAAGCCATATTAACAAGATTGCTCACATTTGTGTGGTGC 856  
Db 826 TCAGCTGTGTCAGCCAGCCAAAAGCCATATTAACAAGATTGCTCACATTTGTGTGGTGC 885  
QY 857 TGAACCGGAGAGATCTATGCACTGCCTGACCCCTACTGTCCCGACAGTGACATCAAAAGC 916  
Db 886 TGAACCGGAGAGATCTATGCACTGCCTGACCCCTACTGTCCCGACAGTGACATCAAAAGC 945

QY 917 CCTCACTACACTGTGTGACTTGGCCGACCGAGAGAGTTGGTGTATTCATTGGATGGCGAA 976  
Db 946 CCTCACTACACTGTGTGACTTGGCCGACCGAGAGAGTTGGTGTATTCATTGGATGGCGAA 1005  
QY 977 GCATATTCCAGGCTTCTCCACGCTGTCCCTGGCGGACCAATGAGCCTTCTGCAAGTGC 1036  
Db 1006 GCATATTCCAGGCTTCTCCACGCTGTCCCTGGCGGACCAATGAGCCTTCTGCAAGTGC 1065  
QY 1037 TTGGATGGAATTTTGAATCCTTGCTGTGCTATACCGGCTCTTTTCATTGAGATGAAC 1096  
Db 1066 TTGGATGGAATTTTGAATCCTTGCTGTGCTATACCGGCTCTTTTCATTGAGATGAAC 1125  
QY 1097 TGTCTATGCAAGCAGATTAATATATGAGCAGAACCACTCCAAATTAGCAGGCTTCTTGA 1156  
Db 1126 TGTCTATGCAAGCAGATTAATATATGAGCAGAACCACTCCAAATTAGCAGGCTTCTTGA 1185  
QY 1157 TCTAATAATGCTATTCCTGCAGCTGGTAAAGAAATACAGACATGAAGCTGAAAAAGA 1216  
Db 1186 TCTAATAATGCTATTCCTGCAGCTGGTAAAGAAATACAGACATGAAGCTGAAAAAGA 1245  
QY 1217 AGAATTTGTACCCCTCAAGCTATAGCTCTTGCTTAATTCAGACTCCATGCACATAGAGA 1276  
Db 1246 AGAATTTGTACCCCTCAAGCTATAGCTCTTGCTTAATTCAGACTCCATGCACATAGAGA 1305  
QY 1277 TGTTGAAGCCGTTACAGAGCTTCAGGATGTCTTACATGAAGCGCTGCAGATTTATGAAGC 1336  
Db 1306 TGTTGAAGCCGTTACAGAGCTTCAGGATGTCTTACATGAAGCGCTGCAGATTTATGAAGC 1365  
QY 1337 TGGCCAGCACATGGAAGACCCCTCGTGCAGCTGGCAAGATGCTGATGACACTGCCACTCT 1396  
Db 1366 TGGCCAGCACATGGAAGACCCCTCGTGCAGCTGGCAAGATGCTGATGACACTGCCACTCT 1425  
QY 1397 GAGGCAAGCCTCTACCAAGGCCGTGCAGCATTTCTAACATCAAACTYAGAAAGCAAAGT 1456  
Db 1426 GAGGCAAGCCTCTACCAAGGCCGTGCAGCATTTCTAACATCAAACTYAGAAAGCAAAGT 1485  
QY 1457 CCCAATGCACAAACTTTTGTGAAATGTTGAGGCCAAGGTCTGACTAAAGCTCCCTG 1516  
Db 1486 CCCAATGCACAAACTTTTGTGAAATGTTGAGGCCAAGGTCTGACTAAAGCTCCCTG 1545  
QY 1517 GGCCCTTCCCATCCTTCATGTTGAAAAAGGAAAAATAAACCAAGAGTGATGCGAAGAA 1576  
Db 1546 GGCCCTTCCCATCCTTCATGTTGAAAAAGGAAAAATAAACCAAGAGTGATGCGAAGAA 1605  
QY 1577 CTTAGAGTTTAACTTAACAATCAAAATCAACAGACTGCACTGATATTTAGCAGCAAG 1636  
Db 1606 CTTAGAGTTTAACTTAACAATCAAAATCAACAGACTGCACTGATATTTAGCAGCAAG 1665  
QY 1637 ACTATGAAGCAGCTTCAGATTCCTCCATAGGTTCCCTGATGAG-TTCTTTCTACTTCTC 1695  
Db 1666 ACTATGAAGCAGCTTCAGATTCCTCCATAGGTTCCCTGATGAGTTTCTTCTACTTCTC 1725  
QY 1696 CATCATCTTCTTCTCTTCTTCTTCCACATTTCTTCTCTTATTTTCTCTTCTTTC 1755  
Db 1726 CATCATCTTCTTCTCTTCTTCTTCCACATTTCTTCTCTTATTTTCTCTTCTTTC 1785  
QY 1756 TTCTTCAACCTCCCTTAATTTCTTGTGCTTCTTCAATTCCTAGTTCOCATTCCTTAAATT 1815  
Db 1786 TTCTTCAACCTCCCTTAATTTCTTGTGCTTCTTCAATTCCTAGTTCOCATTCCTTAAATT 1845  
QY 1816 TCTTCCCGTCTGCGCTTCTTCTTCTTCTTCTTACCTCACTCTCATTCCTCTCTTCTCA 1875  
Db 1846 TCTTCCCGTCTGCGCTTCTTCTTCTTCTTCTTACCTCACTCTCATTCCTCTCTTCTCA 1905  
QY 1876 TCCCTTCCCTTTTCTTAAATTTGAATAGCTTTAGTTTAAAAAATAATCCCTCCCTC 1935  
Db 1906 TCCCTTCCCTTTTCTTAAATTTGAATAGCTTTAGTTT--AAAAAATAATCCCTCCCTC 1963  
QY 1936 CCCCTTTCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1995  
Db 1964 CCCCTTTCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2023

QY 1996 CTGACCTTCTTCCATCTTCTTTCTTCTGCTGCTGAACCTTTTAAAGAGCTC 2055  
|||  
Db 2024 CTGACCTTCTTCCATCTTCTTTCTTCTGCTGCTGAACCTTTTAAAGAGCTC 2083  
QY 2056 TCTAAGAGAGAGATGAGCCAGCCCTGCCAAAGATGAGATCCATAATATGATG 2115  
|||  
Db 2084 TCTAAGAGAGAGATGAGCCAGCCCTGCCAAAGATGAGATCCATAATATGATG 2143  
QY 2116 CCAAGTGAATTTGTGAACCATACCGTCCCAATGACTTAAGGAATCAAGAGAGAGAAC 2175  
|||  
Db 2144 CCAAGTGAATTTGTGAACCATACCGTCCCAATGACTTAAGGAATCAAGAGAGAGAAC 2203  
QY 2176 CAACGTTCTAAAGATGAGTGAACATATACAAATGACTGAGTGAATTTAGATTTC 2235  
|||  
Db 2204 CAACGTTCTAAAGATGAGTGAACATATACAAATGACTGAGTGAATTTAGATTTC 2263  
QY 2236 ATGGAGCAGCCCTTAATTAAGCAACCTTAAGCAACGTTGATCGGCTGCTTATCAT 2295  
|||  
Db 2264 ATGGAGCAGCCCTTAATTAAGCAACCTTAAGCAACGTTGATCGGCTGCTTATCAT 2323  
QY 2296 GCTTTTCCATCTAGATCAGTTACAGCCATTTGATTCCTAATTTGTTTTCAGTCTTC 2355  
|||  
Db 2324 GCTTTTCCATCTAGATCAGTTACAGCCATTTGATTCCTAATTTGTTTTCAGTCTTC 2383  
QY 2356 AGGATTTGTTAGTTAGTCTACTATGTAACCTTTTTCAGGGAATAGTTTAACTTATTC 2415  
|||  
Db 2384 AGGATTTGTTAGTTAGTCTACTATGTAACCTTTTTCAGGGAATAGTTTAACTTATTC 2443  
QY 2416 TTCAATGCAATACTAAGAGAAATAAGAAATGTAATTTGCTGCTTGAACAATTA 2475  
|||  
Db 2444 TTCAATGCAATACTAAGAGAAATAAGAAATGTAATTTGCTGCTTGAACAATTA 2503  
QY 2476 CGAACAATTAATGAAGGACAAATGAATCTCGAAGGAAGATTTTAAATGTTTGTCT 2535  
|||  
Db 2504 CGAACAATTAATGAAGGACAAATGAATCTCGAAGGAAGATTTTAAATGTTTGTCT 2563  
QY 2536 TCTTAACAATGAGATTTTGTGACCAAGCTTTACCACTTTTTCAGCCATTTATTAATG 2595  
|||  
Db 2564 TCTTAACAATGAGATTTTGTGACCAAGCTTTACCACTTTTTCAGCCATTTATTAATG 2623  
QY 2596 GGAATTTAATCTTCAAGCAATAGTTGAAGGGAAGGTGCAATTTATCAGGATGCAAT 2655  
|||  
Db 2624 GGAATTTAATCTTCAAGCAATAGTTGAAGGGAAGGTGCAATTTATCAGGATGCAAT 2683  
QY 2656 TATGTTGTGCGCAGTCTGTCCTCAACATCAATTTCTTAACATGAGCTCCAGTTTACCT 2715  
|||  
Db 2684 TATGTTGTGCGCAGTCTGTCCTCAACATCAATTTCTTAACATGAGCTCCAGTTTACCT 2743  
QY 2716 AATGTTCACTGACACAAAGATGAGATTTACACCTACAGTGAATCTGAGTAGTACATAT 2775  
|||  
Db 2744 AATGTTCACTGACACAAAGATGAGATTTACACCTACAGTGAATCTGAGTAGTACATAT 2803  
QY 2776 ATAAGCACTGCACATGAGATATGATCCGTAGAAATGTCAGAGTGCACCTCTCTACTTG 2835  
|||  
Db 2804 ATAAGCACTGCACATGAGATATGATCCGTAGAAATGTCAGAGTGCACCTCTCTACTTG 2863  
QY 2836 GGAGGTACAATGCGCATATGATTTCTAGCTGCCATGTTGTTAGGAATGTGATACTGCT 2895  
|||  
Db 2864 GGAGGTACAATGCGCATATGATTTCTAGCTGCCATGTTGTTAGGAATGTGATACTGCT 2923  
QY 2896 GTTGGCAAGTCAAGACCTTGCCTGAGAAAGAGCTGTGAGCCAGTATTCATTTAAGAG 2954  
|||  
Db 2924 GTTGGCAAGTCAAGACCTTGCCTGAGAAAGAGCTGTGAGCCAGTATTCATTTAAGAG 2982

RESULT 6  
ADD69756  
ID ADD69756 standard; DNA; 5216 BP.  
XX  
AC ADD69756;  
XX  
DT 15-JAN-2004 (first entry)  
XX

DE Human ERR gamma 3-related DNA - SEQ ID 5.  
XX  
KW nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;  
KW oestrogen receptor; ER; thyroid hormone; TR; human; ds; gene.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 155..1531  
FT /tag= a  
FT /product= "Human ERR gamma 3-related protein - SEQ ID 6"  
XX  
PN WO2003080831-A1.  
XX  
PD 02-OCT-2003.  
XX  
PF 25-MAR-2003; 2003WO-JP003611.  
XX  
PR 25-MAR-2002; 2002JP-00084560.  
XX  
PA (FUJI ) FUJISAWA PHARM CO LTD.  
XX  
PI Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;  
XX  
DR MPI; 2003-779262/73.  
XX  
DR P-PSDB; ADD69757.  
XX  
PT Polynucleotides encoding nuclear receptors, and the encoded proteins,  
PT useful as diagnostic agents, and for identification of agents that affect  
PT receptor activity.  
XX  
PS Claim 1; SEQ ID NO 5; 148bp; Japanese.  
XX  
CC The invention relates to novel nuclear receptor ERR (oestrogen receptor-  
CC related receptor) gamma 3 polynucleotides. The polynucleotides of the  
CC invention may be useful for diagnosis of disorders caused by abnormal  
CC nuclear receptor activity, particularly those related to abnormal  
CC oestrogen receptor (ER), ERR or thyroid hormone receptor (TR) activity.  
CC Furthermore, the polynucleotides and proteins may be useful for  
CC evaluating agents that affect the activity of nuclear receptors. The  
CC current sequence is that of the human ERR gamma 3-related DNA of the  
CC invention.  
XX  
SQ Sequence 5216 BP; 1568 A; 1019 C; 974 G; 1655 T; 0 U; 0 Other;  
QY 17 GTGGAATTCGGCTGTGCTACCTAGAGAAACATTTGTTAATTCAGTGTGCTGTCAAGG 76  
|||  
Db 46 GTGCTATPAGGCTGTGCTACCTAGAGAAACATTTGTTAATTCAGTGTGCTGTCAAGG 105  
QY 77 AAATCTTGATTATAGCTGGGTGCAAAATATGTTGCGGTGCAATGATTCGGT 136  
|||  
Db 106 AAATCTTGATTATAGCTGGGTGCAAAATATGTTGCGGTGCAATGATTCGGT 165  
QY 137 AGAATTTGCTTCTGTAATCTTTTCCCTGCACTAGAGAGAGCTTCTGCAAGAT 196  
|||  
Db 166 AGAATTTGCTTCTGTAATCTTTTCCCTGCACTAGAGAGAGCTTCTGCAAGAT 225  
QY 197 GTCAACAAGATGACACATGATTTCCAGCTGTTGCTCTTCATCAAGAGGGAACCTTC 256  
|||  
Db 226 GTCAACAAGATGACACATGATTTCCAGCTGTTGCTCTTCATCAAGAGGGAACCTTC 285  
QY 257 CAGCCAGCTTCCCTGACGGAGACGCTCAACCAACACAGCCCTGCTTCAAGAGC 316  
|||  
Db 286 CAGCCAGCTTCCCTGACGGAGACGCTCAACCAACACAGCCCTGCTTCAAGAGC 345  
QY 317 CAGTGGAGCTACAGTTCAACCATGATGGCCATCAGAAAGGACTTGACTGCCACCTCT 376  
|||  
Db 346 CAGTGGAGCTACAGTTCAACCATGATGGCCATCAGAAAGGACTTGACTGCCACCTCT 405



QY 377 CTACCCCTTCTGCTCTATCCTGGAGGTAGTGGCCTGTCAGGAAACTGTATGATGACTG 436  
Db 406 CTACCCCTTCTGCTCTATCCTGGAGGTAGTGGCCTGTCAGGAAACTGTATGATGACTG 465  
QY 437 CTCCAGCACCATTGTTGAAGATCCCCAGACCAGTGTGAATACATGCTCACTCGATGCC 496  
Db 466 CTCCAGCACCATTGTTGAAGATCCCCAGACCAGTGTGAATACATGCTCACTCGATGCC 525  
QY 497 CAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACCCTATGGGGTAGCATC 556  
Db 526 CAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACCCTATGGGGTAGCATC 585  
QY 557 ATGTGAAGCCTGCAGGCACTTCTCAAGAGGACAAATTCAGGCAATATAGATACAGCTG 616  
Db 586 ATGTGAAGCCTGCAGGCACTTCTCAAGAGGACAAATTCAGGCAATATAGATACAGCTG 645  
QY 617 CCCTGCCAGCAATGAATGTGAATCACAAGCGGACGTAATCCTGCCAGGCTTGCCG 676  
Db 646 CCCTGCCAGCAATGAATGTGAATCACAAGCGGACGTAATCCTGCCAGGCTTGCCG 705  
QY 677 CTTCATGAAGTGTTTAAAGTGGGCATGCTGAAGAAGGGGTGCGTCTTGACAGAGTACG 736  
Db 706 CTTCATGAAGTGTTTAAAGTGGGCATGCTGAAGAAGGGGTGCGTCTTGACAGAGTACG 765  
QY 737 TGGAGGTCGGCAGAAATCAAGCCGAGATTAGATGCGGAGAAACAGCCCATCCTGAACCC 796  
Db 766 TGGAGGTCGGCAGAAATCAAGCCGAGATTAGATGCGGAGAAACAGCCCATCCTGAACCC 825  
QY 797 TCAGCTGGTTCAGCCAGCCAAAAAGCCATATTAACAAGATTGTCTACATTTGTTGGTGC 856  
Db 826 TCAGCTGGTTCAGCCAGCCAAAAAGCCATATTAACAAGATTGTCTACATTTGTTGGTGC 885  
QY 857 TGAACCGGAGAAGATCTATGCCATGCGCTGACCCCTACTGTCCCCGACAGTGAATCAAAAGC 916  
Db 886 TGAACCGGAGAAGATCTATGCCATGCGCTGACCCCTACTGTCCCCGACAGTGAATCAAAAGC 945  
QY 917 CCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGGTGTATCATTTGAGGGCGAA 976  
Db 946 CCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGGTGTATCATTTGAGGGCGAA 1005  
QY 977 GCATATTTCCAGGCTTTCACACGCTGTCCCTGGCGGACCAAGTAGCCTTTCGACAGATGC 1036  
Db 1006 GCATATTTCCAGGCTTTCACACGCTGTCCCTGGCGGACCAAGTAGCCTTTCGACAGATGC 1065  
QY 1037 TTGGATGGAATTTTGATCTTGGTGTGCTGATACCGGTCTCTTCATTTGAGATGAACCT 1096  
Db 1066 TTGGATGGAATTTTGATCTTGGTGTGCTGATACCGGTCTCTTCGTTGAGATGAACCT 1125  
QY 1097 TGTCTATGCAGACGATTAATAATGAGCGAAGACCAATTCAGAGCCCTTCTTGA 1156  
Db 1126 TGTCTATGCAGACGATTAATAATGAGCGAAGACCAATTCAGAGCCCTTCTTGA 1185  
QY 1157 TCTAATAATGCTATCTGCAGCTGTGTAAGAATAACAAGCATGAAAGCTGAAAAAGA 1216  
Db 1186 TCTAATAATGCTATCTGCAGCTGTGTAAGAATAACAAGCATGAAAGCTGAAAAAGA 1245  
QY 1217 AGAATTTGTCAACCCTCAAAGCTATAGCTCTTGCTAATTCAAGCTCCATGCAATGAAGA 1276  
Db 1246 AGAATTTGTCAACCCTCAAAGCTATAGCTCTTGCTAATTCAAGCTCCATGCAATGAAGA 1305  
QY 1277 TGTGTAAGCCGTTCAAGAGCTTCAGGATGTCTTACATGAAGCGCTGCAGGATTAATGAAGC 1336  
Db 1306 TGTGTAAGCCGTTCAAGAGCTTCAGGATGTCTTACATGAAGCGCTGCAGGATTAATGAAGC 1365  
QY 1337 TGGCCAGACATGGAAGACCTCTGTCAGCTGGCAAGATGCTGATGACACTGCCACTCCT 1396  
Db 1366 TGGCCAGACATGGAAGACCTCTGTCAGCTGGCAAGATGCTGATGACACTGCCACTCCT 1425  
QY 1397 GAGGCAAGACCTCTACCAAGCCGTGCAGCATTTCTACAACATCAAACTAGAAGGCAAGT 1456  
Db 1426 GAGGCAAGACCTCTACCAAGCCGTGCAGCATTTCTACAACATCAAACTAGAAGGCAAGT 1485  
QY 1457 CCCAATGCACAAACTTTTGTGAAATGTGTGAGGCCAAGGTCTGACTAAAAAGCTCCCTG 1516

Db 1486 CCCAATGCACAAACTTTTGTGAAATGTGAGGCCAAGGTCTGACTAAAAAGCTCCCTG 1545  
QY 1517 GGCCTTCCCATCTTCATGTTGAAAAAGGAAAAATAAACCCAGAGTGTGGAAGAAA 1576  
Db 1546 GGCCTTCCCATCTTCATGTTGAAAAAGGAAAAATAAACCCAGAGTGTGGAAGAAA 1605  
QY 1577 CTTAGAGTTTACTTAACAACATCAAAAAATCAACAGACTGCATGATTAATTAGCAGCAAG 1636  
Db 1606 CTTAGAGTTTACTTAACAACATCAAAAAATCAACAGACTGCATGATTAATTAGCAGCAAG 1665  
QY 1637 ACTATGAAGCAGCTTTCAGATTCCCTCATAGTTCCTGATGAG-TTCTTTCTACTTCTC 1695  
Db 1666 ACTATGAAGCAGCTTTCAGATTCCCTCATAGTTCCTGATGAGTTCTTTCTACTTCTC 1725  
QY 1696 CATCATCTTTCTCTCTTTCTTCCACATTTCTTTCTCTTATTTTCTCTCTTTC 1755  
Db 1726 CATCATCTTTCTCTCTTTCTTCCACATTTCTTTCTCTTATTTTCTCTCTTTC 1785  
QY 1756 TTTCTTCAACCTCCCTTAATTTCTTGTCTTCTTCAATTCCTAGTTCCTCATTTCTTAATT 1815  
Db 1786 TTTCTTCAACCTCCCTTAATTTCTTGTCTTCTTCAATTCCTAGTTCCTCATTTCTTAATT 1845  
QY 1816 TCTTCCCGTCTGCGCTGCTTCTTCTTCTTCTTCACTCTCATTCCTCTCTTCTCA 1875  
Db 1846 TCTTCCCGTCTGCGCTGCTTCTTCTTCTTCTTCACTCTCATTCCTCTCTTCTCA 1905  
QY 1876 TCCTTCCCGTTTTTCTAAATTTGAATAGCTTAAGTTAAAAAATAATCTCCCTTC 1935  
Db 1906 TCCTTCCCGTTTTTCTAAATTTGAATAGCTTAAGTT-AAAAAATAATCTCCCTTC 1963  
QY 1936 CCCCTTTCCTTTCCTTCTTCTTCTTTCCTTTCCTTTCCTTTCCTTTCCT 1995  
Db 1964 CCCCTTTCCTTTCCTTCTTCTTCTTTCCTTTCCTTTCCTTTCCTTTCCT 2023  
QY 1996 CTTGACCTTCTTTCATCTTTCTTTCTTCTTCTGCTGCTGAACCTTTTAAAGAGTTC 2055  
Db 2024 CTTGACCTTCTTTCATCTTTCTTTCTTCTTCTGCTGCTGAACCTTTTAAAGAGTTC 2083  
QY 2056 TCTAATGAAGAGATGGAAGCCAGCCCTGCCAAAGAGATGAGATCCATAATATGATG 2115  
Db 2084 TCTAATGAAGAGATGGAAGCCAGCCCTGCCAAAGAGATGAGATCCATAATATGATG 2143  
QY 2116 CCAATGAACCTTAATGGAACCATATACCGTCCCAATGACTAAGGAATCAAGAGAGAGAAC 2175  
Db 2144 CCAATGAACCTTAATGGAACCATATGCTCCCAATGACTAAGGAATCAAGAGAGAGAAC 2203  
QY 2176 CAACGTTCCCTAAAGATACAGTGCACATATACAATGACTGAGTGCAGTATTAATTC 2235  
Db 2204 CAACGTTCCCTAAAGATACAGTGCACATATACAATGACTGAGTGCAGTATTAATTC 2263  
QY 2236 ATGGGAGCAGCCTCTAATTAAGCAACTTAAGCAACGTTGCATCGGCTGCTTTATCATTT 2295  
Db 2264 ATGGGAGCAGCCTCTAATTAAGCAACTTAAGCAACGTTGCATCGGCTGCTTTATCATTT 2323  
QY 2296 GCTTTTCCATCTAGATCAGTTACAGCCATTTGATTCCTTAATGTTTTTCAAGTCTTC 2355  
Db 2324 GCTTTTCCATCTAGATCAGTTACAGCCATTTGATTCCTTAATGTTTTTCAAGTCTTC 2383  
QY 2356 AGGTATTTGTTAGTTTAACTACTATGTAACCTTTTCAAGGAATAGTTTAAAGCTTAATCA 2415  
Db 2384 AGGTATTTGTTAGTTTAACTACTATGTAACCTTTTCAAGGAATAGTTTAAAGCTTAATCA 2443  
QY 2416 TTCAATGCAATACTAAGAGCAAAATTAAGATACGCAATTTTGTGCTGGCTTGAACAAATTA 2475  
Db 2444 TTCAATGCAATACTAAGAGCAAAATTAAGATACGCAATTTTGTGCTGGCTTGAACAAATTA 2503  
QY 2476 CGAACATATATGAAGCAAAATGAATCCTGAAGGAAGATTTTAAAAATGTTTGTTCCT 2535  
Db 2504 CGAACATATATGAAGCAAAATGAATCCTGAAGGAAGATTTTAAAAATGTTTGTTCCT 2563  
QY 2536 TCTTAAATGAGATTTTTTTGTACCAGCTTTACCACTTTTCAAGCATTTATTAATATG 2595



Db 2564 TCTTACAATGAGATTTTTTTGTACACGCTTTACCACTTTTCAGCCATTATTAATATG 2623  
QY 2596 GGAATTTAACTTACTCAAGAAATAGTTGAAGGGAAGTGCATATTATCACGGATGCAATT 2655  
Db 2624 GGAATTTAACTTACTCAAGCAATAGTTGAAGGGAAGTGCATATTATCACGGATGCAATT 2683  
QY 2656 TATGTTGTCGACGTCGTGTCGCCAACAACATCAATTTCTTAACATGAGCTCCAGTTTACCT 2715  
Db 2684 TATGTTGTCGACGTCGTGTCGCCAACAACATCAATTTCTTAACATGAGCTCCAGTTTACCT 2743  
QY 2716 AAATGTTCACTGCACACAAGATGAGATTACACCTTACAGTGACTCTGAGTAGTCAATAT 2775  
Db 2744 AAATGTTCACTGCACACAAGATGAGATTACACCTTACAGTGACTCTGAGTAGTCAATAT 2803  
QY 2776 ATAAGCACTGCACATGAGATATAGATCCGTAGATTGTCAAGAGTGCACCTCTTACTTG 2835  
Db 2804 ATAAGCACTGCACATGAGATATAGATCCGTAGATTGTCAAGAGTGCACCTCTTACTTG 2863  
QY 2836 GGAGGTACAATTGCCATATGATTTCTAGCTGCCATGCTGTTAGGAATGTGATACTGCCCT 2895  
Db 2864 GGAGGTACAATTGCCATATGATTTCTAGCTGCCATGCTGTTAGGAATGTGATACTGCCCT 2923  
QY 2896 GTTTCAAAAGTCAACAGACCTTGCTCAGAAAGAGCTGTGAGCCAGTATTCATTTAAGAG 2954  
Db 2924 GTTTCAAAAGTCAACAGACCTTGCTCAGAAAGAGCTGTGAGCCAGTATTCATTTAAGAG 2982

RESULT 7  
ADD69754  
ID ADD69754 standard; DNA; 5221 BP.  
XX  
AC ADD69754;  
XX  
DT 15-JAN-2004 (first entry)

XX Human ERR gamma 3-related DNA - SEQ ID 3.  
XX  
KW nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;  
KM oestrogen receptor; ER; thyroid hormone; TR; human; ds; gene.  
XX Homo sapiens.

XX  
FH Key Location/Qualifiers  
FT CDS 219..1529  
FT /\*tag= a  
FT /product= "Human ERR gamma 3-related protein - SEQ ID 4"

XX  
PN WO2003080831-A1.  
XX  
PD 02-OCT-2003.  
XX  
PF 25-MAR-2003; 2003WO-JP003611.  
XX  
PR 25-MAR-2002; 2002JP-00084560.  
XX

XX  
PA (FUJI ) FUJISAWA PHARM CO LTD.  
XX  
PI Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;  
XX  
DR WPI: 2003-779262/73.  
DR P-PSDB; ADD69755.

XX  
PT Polynucleotides encoding nuclear receptors, and the encoded proteins,  
PT useful as diagnostic agents, and for identification of agents that affect  
PT receptor activity.

XX  
PS Claim 9; SEQ ID NO 3; 148bp; Japanese.

XX  
CC The invention relates to novel nuclear receptor ERR (oestrogen receptor-  
CC related receptor) gamma 3 polynucleotides. The polynucleotides of the  
CC invention may be useful for diagnosis of disorders caused by abnormal  
CC nuclear receptor activity, particularly those related to abnormal  
CC oestrogen receptor (ER), ERR or thyroid hormone receptor (TR) activity.

CC Furthermore, the polynucleotides and proteins may be useful for  
CC evaluating agents that affect the activity of nuclear receptors. The  
CC current sequence is that of the human ERR gamma 3-related DNA of the  
CC invention.

SQ Sequence 5221 BP; 1584 A; 1037 C; 967 G; 1633 T; 0 U; 0 Other;

Query Match 88.0%; Score 2626; DB 9; Length 5221;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 2738; Conservative 0; Mismatches 35; Indels 11; Gaps 8;

QY 177 GAAGAGCTTCTGCAAGATGTCAAAACAAGATGCACATTTGATTCAGCTGTGTC 236  
Db 201 GAAAGCTTCTGCAAGATGTCAAAACAAGATGCACATTTGATTCAGCTGTGTC 260  
QY 237 TTCATCAAGACGGAACCTTCCAGCCCAAGCTCCCTGACCGACAGCGTCAACCAACAGC 296  
Db 261 TTCATCAAGACGGAACCTTCCAGCCCAAGCTCCCTGACCGACAGCGTCAACCAACAGC 320  
QY 297 CCTGGTGGCTCTTCAAGACCCAGTGGAGCTTACAGTTCAACCATGAATGGCCATCAGAAC 356  
Db 321 CCTGGTGGCTCTTCAAGACCCAGTGGAGCTTACAGTTCAACCATGAATGGCCATCAGAAC 380  
QY 357 GGAATTGACTGGCCACTCTCTACCCCTTGTCTATCTCTGGAGGTAGTGGCTGTC 416  
Db 381 GGAATTGACTGGCCACTCTCTACCCCTTGTCTATCTCTGGAGGTAGTGGCTGTC 440  
QY 417 AGGAAACTGTATGATGACTGCTCCAGCACCATTGTTGAAGATCCCAGACCAAGTGTGA 476  
Db 441 AGGAAACTGTATGATGACTGCTCCAGCACCATTGTTGAAGATCCCAGACCAAGTGTGA 500  
QY 477 TACATGCTCAACTCGATGATGCCCAAGAGACTGTGTTAGTGTGTGATGACATGCTTGGG 536  
Db 501 TACATGCTCAACTCGATGATGCCCAAGAGACTGTGTTAGTGTGTGATGACATGCTTGGG 560  
QY 537 TACCACCTATGGGTAGCATCATGTGAAGCTGCAAGGCATTCTCAAGAGACAATTCAA 596  
Db 561 TACCACCTATGGGTAGCATCATGTGAAGCTGCAAGGCATTCTCAAGAGACAATACAA 620  
QY 597 GGCAATATAGAAATACAGCTGCCCTGCCCGAATGAATGTGAATCACAAAGCCGACAGCT 656  
Db 621 GGCAATATAGAAATACAGCTGCCCTGCCCGAATGAATGTGAATCACAAAGCCGACAGCT 680  
QY 657 AAATCTGCGAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGATGCTGAAGAAGGG 716  
Db 681 AAATCTGCGAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGATGCTGAAGAAGGG 740  
QY 717 GTGCGTCTTGACAGAGTACGTGAGGTGCGCAGAAATCAAGCGCAGATAGATGGGAG 776  
Db 741 GTGCGTCTTGACAGAGTACGTGAGGTGCGCAGAAATCAAGCGCAGATAGATGGGAG 800  
QY 777 AACAGCCCATACCTGAACCTCAGCTGTTCAAGCCCAAAAAGCCATATTAACAAGATT 836  
Db 801 AACAGCCCATACCTGAACCTCAGCTGTTCAAGCCCAAAAAGCCATATTAACAAGATT 860  
QY 837 GTCTCACATTTGTTGGTGGCTGAACCGGAGAAATCTATGCCATGCTGACCCTACTGTC 896  
Db 861 GTCTCACATTTGTTGGTGGCTGAACCGGAGAAATCTATGCCATGCTGACCCTACTGTC 920  
QY 897 CCCGACAGTGACATCAAAAGCCCTCACTACACTGTGTGACTTGCGCCGACCGAGAGTTGGTG 956  
Db 921 CCCGACAGTGACATCAAAAGCCCTCACTACACTGTGTGACTTGCGCCGACCGAGAGTTGGTG 980  
QY 957 GTTATCATTTGATGGGGAAGCATATTTCCAGGCTTCTCCAGCTGTGCGGGAACAG 1016  
Db 981 GTTATCATTTGATGGGGAAGCATATTTCCAGGCTTCTCCAGCTGTGCGGGAACAG 1040  
QY 1017 ATGAGCCTTCTGACAGAGTCTTGATGAAATTTGATCCTTGTGTGATACCGGTCT 1076  
Db 1041 ATGAGCCTTCTGACAGAGTCTTGATGAAATTTGATCCTTGTGTGATACCGGTCT 1100  
QY 1077 CTTTCATTTGAGAGTGAAGTGTCTATGACAGAGATATATATATGACGAAGACAGTCC 1136

```
Db 1101 CTTTCGTTTGAGGATGAACCTGTCTATGCAAGCATTTATATATGACGAAGACCAGTCC 1160
QY 1137 AAATTAGCAGGCTTCTTGATCTAAATATATGCTATCCTGCAGCTGGTAAAGAAATACAAG 1196
Db 1161 AAATTAGCAGGCTTCTTGATCTAAATATATGCTATCCTGCAGCTGGTAAAGAAATACAAG 1220
QY 1197 AGCATGAAGCTGSAAGAAAGAAATTTGTGACCCCTCAAGCTATAGCTCTTGCTAATTCA 1256
Db 1221 AGCATGAAGCTGSAAGAAAGAAATTTGTGACCCCTCAAGCTATAGCTCTTGCTAATTCA 1280
QY 1257 GACTCCATGACATAGAAAGATGTTGAAGCCGTTCAAGACTTCAGAGTGTCTTACATGAA 1316
Db 1281 GACTCCATGACATAGAAAGATGTTGAAGCCGTTCAAGACTTCAGAGTGTCTTACATGAA 1340
QY 1317 GCGCTGACAGATTATGAAGCTGGCCAGACATGGAAGACCCCTGTCGAGCTGGCAAGATG 1376
Db 1341 GCGCTGACAGATTATGAAGCTGGCCAGACATGGAAGACCCCTGTCGAGCTGGCAAGATG 1400
QY 1377 CTGATGACACTGCCACTCCTCGAGCAGACCTCTACCAAGCCGTGACGATTTCTACAAC 1436
Db 1401 CTGATGACACTGCCACTCCTCGAGCAGACCTCTACCAAGCCGTGACGATTTCTACAAC 1460
QY 1437 ATCAAACTAGAAAGCCAAAGTCCCAATGCACAACTTTTTTGGAAATGTTGAGGCCAAG 1496
Db 1461 ATCAAACTAGAAAGCCAAAGTCCCAATGCACAACTTTTTTGGAAATGTTGAGGCCAAG 1520
QY 1497 GTCTGACTAAAGCTCCCTGGGCCCTCCCAATCTTCATGTTGAAAAAGGGAATTAACC 1556
Db 1521 GTCTG-CTAAAGCTCCCTGGGCCCT-CCATCTTCAITGTTGAAAAAGGGAATTAACC 1578
QY 1557 CAAGAGTGATGTCGAGAAGAACTTAGAGTTTAACTAACACATCAAAAAATCAACAGACTGC 1616
Db 1579 CAAGAGTGATGTCGAGAAGAACTTAGAGTTTAACTAACACATCAAAAAATCAACAGACTGC 1638
QY 1617 ACTGATAATTAGCAGCAAGACTATGAAAGCAGCTTTCAGATTCCTCCATAGTTCTCTGAT 1676
Db 1639 ACTGATAATTAGCAGCAAGACTATGAAAGCAGCTTTCAGATTCCTCCATAGTTCTCTGAT 1698
QY 1677 GAG-TTCTTTCTACTTCTCCATCATCTTCTCTCTCTTCTCCACATTTCTCTTCT 1735
Db 1699 GAGTTTCTTCTACTTCTCCATCATCTTCTCTCTCTTCTCCACATTTCTCTTCT 1758
QY 1736 CTTTATTTTTCTCTTTTCTTCTTTCACCTCCCTTAATTTCTTGTGCTTCTTCAATCCTA 1795
Db 1759 CTTTATTTTTATCTCTTCTTCTTTCACCTCCCTTAATTTCTTGTGCTTCTTCAATCCTA 1818
QY 1796 GTTCCCATCTCTCTTAATTTTCTTCCGCTGCTGCTCTTCTTCTTTTCACTACT 1855
Db 1819 GTTCCCATCTCTCTTAATTTTCTTCCGCTGCTGCTCTTCTTCTTTTCACTACT 1878
QY 1856 CTGATTCCTCTCTTTTCTCATCTTCCCTTTTCTAAATTTGAATATAGCTTAGTTTA 1915
Db 1879 CTGATTCCTCTCTTTTCTCATCTTCCCTTTTCTAAATTTGAATATAGCTTAGTTT- 1937
QY 1916 AAAAAAAAAATCTCCCTCCCTTCCCTTCTCTTCTCTTCTCTTCTCTTCTCTTCT 1975
Db 1938 -AAAAAAAAATCTCCCTCCCTTCCCTTCTCTTCTCTTCTCTTCTCTTCTCTTCT 1996
QY 1976 TCCCTTCTCTTCTCTTCTCTTGACCTTCTTCCATCTTCTTCTTCTTCTCTCTGCTG 2035
Db 1997 TCCCTTCTCTTCTCTTCTCTTGACCTTCTTCCATCTTCTTCTTCTTCTCTCTGCTG 2056
QY 2036 CTGAACCTTTTAAAGAGTCTCTAATCTGAAGAGAGATGGAAGCCAGCCCTGCCAAAGGAT 2095
Db 2057 CTGAACCTTTTAAAGAGTCTCTAATCTGAAGAGAGATGGAAGCCAGCCCTGCCAAAGGAT 2116
QY 2096 GGAGATCCATATATGATGATGCCAGTGAACCTTATGTGAACCATACCGTCCCAATGACTA 2155
Db 2117 GGAGATCCATATATGATGATGCCAGTGAACCTTATGTGAACCATACCGTCCCAATGACTA 2176
QY 2156 AGGAATCAAAAGAGAGAACCAAGCTTCTTAAAGTACAGTGCACATATATACAAATTGAC 2215
Db 2177 AGGAATCAAAAGAGAGAACCAAGCTTCTTAAAGTACAGTGC-ACATATACAAATTGAC 2235
```

```
QY 2216 TGAGTGAGTATTAGATTTTCATGGGAGCAGCCCTTAATTAGACAACTTAAAGCAAGTTGC 2275
Db 2236 TGAGTGAGTATTAGATTTTCATGGGAGCAGCCCTTAATTAGACAACTTAAAGCAAGTTGC 2295
QY 2276 ATCGCTGCTCTTATCATTTGCTTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTA 2335
Db 2296 ATCGCTGCTCTTATCATTTGCTTTTCCATCTAGAGCAGTTACAGCCATTGATTCCTTA 2355
QY 2336 ATGTTTTTCAAGCTCTTCCAGGTATTTGTTAGTTTACTACTATAGTAACTTTTCAGGG 2395
Db 2356 ATGTTTTTCAAGCTCTCCAGGTATTTGTTAGTTTACTACTATAGTAACTTTTCAGGG 2415
QY 2396 AATAGTTTAACTTTATTCATTCATGCAATTAATAAGAGAAATACTGCAATTTT 2455
Db 2416 AATAGTTTAACTTTATTCATTCATGCAATTAATAAGAGAAATACTGCAATTTT 2475
QY 2456 GTGCTGCTTTGAACAATTAAGCAACAATTAAGGACAAATGAATCCTGAAGGAAGATT 2515
Db 2476 GTGCTGCTTTGAACAATTAAGCAACAATTAAGGACAAATGAATCCTGAAGGAAGATT 2535
QY 2516 TTTAAAAATGTTTGTCTTCTTCTTCAAAATGAGATTTTGTGTACCAGCTTTACCACTT 2575
Db 2536 TTTAAAAATGTTTGTCTTCTTCTTCAAAATGAGATTTTGTGTACCAGCTTTACCACTT 2595
QY 2576 TTCAGCCATTTATTAATATGGAATTAACTTACTCAAGCAATATGTTGAAGGAAGGTGC 2635
Db 2596 TTCAGCCATTTATTAATATGGAATTAACTTACTCAAGCAATATGTTGAAGGAAGGTGC 2655
QY 2636 ATATTATCAGGATGCAATTTATG--TTGTGTGCCAGTCTGTGCCCAATCAATTTTC 2692
Db 2656 ATATTATCAGGATGCAATTTATGTTGTGTGCCAGTCTGTGCCCAATCAATTTTC 2715
QY 2693 TTACATGAGCTCCAGTTTACCTAAATGTTCACTGACACAAAGATGAGATTACACCTAC 2752
Db 2716 TTACATGAGCTCCAGTTTACCTAAATGTTCACTGACACAAAGATGAGATTACACCTAC 2775
QY 2753 AGTACTCTGAGTAGTCACATATATATAGCACTGCAATGAGATATAGATCCGTAGAATTG 2812
Db 2776 AGTACTCTGAGTAGTCACATATATATAGCACTGCAATGAGATATAGATCCGTAGAATTG 2835
QY 2813 TCAGAGTGCACCTCTCTACTTGGGAGGTACAAATGCCATATGATTTCTAGCTGCCATGG 2872
Db 2836 TCAGAGTGCACCTCTCTACTTGGGAGGTACAAATGCCATATGATTTCTAGCTGCCATGG 2895
QY 2873 TGGTTAGGAATGTGATAC-TGCCCTGTTTGCAAGTCAAGACC-TTGCTCAGAAAGAGC 2930
Db 2896 TGGTTAGGAATGTGATACATGCTGTTTGCAAGTCAAGACCATTGCTCAGAAAGAGC 2955
QY 2931 TGTGAGCCAGTATTCATTTAAGAG 2954
Db 2956 TGTGAGCCAGTATTCATTTAAGAG 2979

RESULT 8
ADD69752 ID ADD69752 standard; cDNA; 3362 BP.
XX
AC ADD69752;
XX
DE Human ERR gamma 3-related cDNA - SEQ ID 1.
XX
KW nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;
XX oestrogen receptor; ER; thyroid hormone; TR; human; ss; gene.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 684..1874
FT FT /*tag= a
FT FT /product= "Human ERR gamma 3-related protein - SEQ ID 2"
```

XX WO2003080831-A1.  
XX 02-OCT-2003.  
XX 25-MAR-2003; 2003WO-JP003611.  
XX 25-MAR-2002; 2002JP-00084560.  
XX (FUJI ) FUJISAWA PHARM CO LTD.  
XX Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;  
XX WPI; 2003-779262/73.  
XX P-PSDB; ADD69753.  
XX Polynucleotides encoding nuclear receptors, and the encoded proteins,  
XX useful as diagnostic agents, and for identification of agents that affect  
XX receptor activity.  
XX Claim 1; SEQ ID NO 1; 148bp; Japanese.  
XX  
XX The invention relates to novel nuclear receptor ERR (oestrogen receptor-  
XX related receptor) gamma 3 polynucleotides. The polynucleotides of the  
XX invention may be useful for diagnosis of disorders caused by abnormal  
XX nuclear receptor activity, particularly those related to abnormal  
XX oestrogen receptor (ER), ER or thyroid hormone receptor (TR) activity.  
XX Furthermore, the polynucleotides and proteins may be useful for  
XX evaluating agents that affect the activity of nuclear receptors. The  
XX current sequence is that of the human ERR gamma 3-related cDNA of the  
XX invention.  
XX  
XX Sequence 3362 BP; 937 A; 767 C; 667 G; 991 T; 0 U; 0 Other;

Query Match 83.7%; Score 2497.8; DB 9; Length 3362;  
Best Local Similarity 95.1%; Pred. No. 0;  
Matches 2658; Conservative 0; Mismatches 17; Indels 120; Gaps 3;

QY 161 TTCCCTGCACCTACGAGGAGAGCTTCTCTGCAAGATGTCAAACAAGATGCACACATTGA 220  
Db 650 TCCACTGAGAAAGGAGATTAAGGCTTCTCTGCAAGATGTCAAACAAGATGCACACATTGA 709  
QY 221 TTCACGCTGTTCCTCTTCATCAAGACGGAACCTTCCAGCCCGCTCCCTGACGACAG 280  
Db 710 TTCACGCTGTTCCTCTTCATCAAGACGGAACCTTCCAGCCCGCTCCCTGACGACAG 769  
QY 281 CGTCAACCCACCAAGCCTGTGGCTCTTCAAGCCAGTGGAGCTACAGTTCAACCAT 340  
Db 770 CGTCAACCCACCAAGCCTGTGGCTCTTCAAGCCAGTGGAGCTACAGTTCAACCAT 829  
QY 341 GAATGGCCATCAGAACGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 400  
Db 830 GAATGGCCATCAGAACGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 889  
QY 401 AGGTAGTGGGCTGTCAAGAACTGTATGATGACTGCTCCAGCACCATTGTTGAAGATCC 460  
Db 890 AGGTAGTGGGCTGTCAAGAACTGTATGATGACTGCTCCAGCACCATTGTTGAAGATCC 949  
QY 461 CCAGACCAAGTGTGAATACATGCTCAACTGCATGCCCAAGAGACTGTGTTAGTGTGG 520  
Db 950 CCAGACCAAGTGTGAATACATGCTCAACTGCATGCCCAAGAGACTGTGTTAGTGTGG 1009  
QY 521 TGACATCGCTTGGGTACCACTATGGGGTACATGTAAGCCTGCAAGGACATTCTT 580  
Db 1010 TGACATCGCTTGGGTACCACTATGGGGTACATGTAAGCCTGCAAGGACATTCTT 1069  
QY 581 CAAGAGACAATTCAAGCAATATAGAAATACAGCTGCCCTGCCAGAAATGAATGTGAAT 640  
Db 1070 CAAGAGACAATTCAAGCAATATAGAAATACAGCTGCCCTGCCAGAAATGAATGTGAAT 1083  
QY 641 CACAAGCGCAGACGTAATCTCGCAGGCTTCCGCTTCATGAAGTGTAAAGTGGG 700  
Db 1084 ----- 1083

QY 701 CATGCTGAAAAGGGGTGCGTCTTGACAGAGTACGTGAGGTCGGCAGAAATACAAAGCG 760  
Db 1084 -----AAGGGGTGCGTCTTGACAGAGTACGTGAGGTCGGCAGAAATACAAAGCG 1132  
QY 761 CAGGATGATGCGGAGAACAGCCCATACCTGAACCTCAGCTGTTACGCCAGCAAAAA 820  
Db 1133 CAGGATGATGCGGAGAACAGCCCATACCTGAACCTCAGCTGTTACGCCAGCAAAAA 1192  
QY 821 GCCATATAACAAGATTGCTCAATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 880  
Db 1193 GCCATATAACAAGATTGCTCAATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1252  
QY 881 GCCTGACCCCTACTGTCCCGACAGTGAATCAAGCCCTCACTACACTGTGTGACTTGGC 940  
Db 1253 GCCTGACCCCTACTGTCCCGACAGTGAATCAAGCCCTCACTACACTGTGTGACTTGGC 1312  
QY 941 CGACCGAGATTGGTGTATTCATTGGATGGCGGAGCATATTCAGGCTTCTCCAGCCT 1000  
Db 1313 CGACCGAGATTGGTGTATTCATTGGATGGCGGAGCATATTCAGGCTTCTCCAGCCT 1372  
QY 1001 GTCCCTGGGGGACGACATGAGCCTTCTGACAGATGCTGGATGGAATTTTGATCCTGG 1060  
Db 1373 GTCCCTGGGGGACGACATGAGCCTTCTGACAGATGCTGGATGGAATTTTGATCCTGG 1432  
QY 1061 TGTGTATACCGGTCTCTTTCATTGAGATGAACCTGTCTATGACAGACGATTAATAAT 1120  
Db 1433 TGTGTATACCGGTCTCTTTCATTGAGATGAACCTGTCTATGACAGACGATTAATAAT 1492  
QY 1121 GGAAGAGACCAAGTCCAAATTAGCAGGCTTCTGATCTAATAATGCTATCCTGAGCT 1180  
Db 1493 GGAAGAGACCAAGTCCAAATTAGCAGGCTTCTGATCTAATAATGCTATCCTGAGCT 1552  
QY 1181 GGTAAAGAATATACAGACATGAGCTGGAAGAAAGAAATTTGTCACCCTCAAGCTAT 1240  
Db 1553 GGTAAAGAATATACAGACATGAGCTGGAAGAAAGAAATTTGTCACCCTCAAGCTAT 1612  
QY 1241 AGCTCTGCTAATTCAGACTCCATGCAATAGAAAGTGTGAAGCCGTTCAAGACTTCA 1300  
Db 1613 AGCTCTGCTAATTCAGACTCCATGCAATAGAAAGTGTGAAGCCGTTCAAGACTTCA 1672  
QY 1301 GGATGCTTACATGAAAGCGCTGCAAGATTATGAAGCTGCCAGCAGATGGAAGCCCTCG 1360  
Db 1673 GGATGCTTACATGAAAGCGCTGCAAGATTATGAAGCTGCCAGCAGATGGAAGCCCTCG 1732  
QY 1361 TCGAGCTGGCAAGATGCTGATGACACTGCACTCTGAGGAGAGACCTCTACAGAGCCGT 1420  
Db 1733 TCGAGCTGGCAAGATGCTGATGACACTGCACTCTGAGGAGAGACCTCTACAGAGCCGT 1792  
QY 1421 GCAGCATTTCTCAACATCAAACTAGAAGGCAAGTCCCAATGCACAAACTTTTTTGA 1480  
Db 1793 GCAGCATTTCTCAACATCAAACTAGAAGGCAAGTCCCAATGCACAAACTTTTTTGA 1852  
QY 1481 AATGTTGAGGCGCAAGGTCTGACTAAAGCTCCCTGGGCTTCCCATCCTTCATGTTGA 1540  
Db 1853 AATGTTGAGGCGCAAGGTCTGACTAAAGCTCCCTGGGCTTCCCATCCTTCATGTTGA 1912  
QY 1541 AAAGGAAATTAACCCCAAGAGTGTGCGAAGAACTTGAAGTTAGTTAACAACATCA 1600  
Db 1913 AAAGGAAATTAACCCCAAGAGTGTGCGAAGAACTTGAAGTTAGTTAACAACATCA 1972  
QY 1601 AAATCAACAGACTGCACTGATAATTATAGAGCAAGACTATGAAGCAGCTTTCAGATTCC 1660  
Db 1973 AAATCAACAGACTGCACTGATAATTATAGAGCAAGACTATGAAGCAGCTTTCAGATTCC 2032  
QY 1661 TCCATAGGTTCTGATGAG-TTCTTTCTACTTCTCCATCATCTTCTTCTTCTTCTTCTC 1719  
Db 2033 TCCATAGGTTCTGATGAGTTCTTCTTCTACTTCTCCATCATCTTCTTCTTCTTCTC 2092  
QY 1720 CCACATTTCTCTTCTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1779  
Db 2093 CCACATTTCTCTTCTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2152



QY	1780	GCTTCTTTCAATTCCTAGTTCCTCCCATTCCTCCTTTAATTTTCTTCCCGTCTGCCCTTCTTT	1839
Db	2153	GCTTCTTTCAATTCCTAGTTCCTCCCATTCCTCCTTTAATTTTCTTCCCGTCTGCCCTTCTTT	2212
QY	1840	CTTTTCTTTAACCCTACTCTCAATTCCTCTCTTTCTCATCCCTTCCCTTTTCTTAAATTTG	1899
Db	2213	CTTTTCTTTAACCCTACTCTCAATTCCTCTCTTTCTCATCCCTTCCCTTTTCTTAAATTTG	2272
QY	1900	AAATAGCTTTAGTTTAAAAAATCCTCCCTTCCCTTCCCTTCCCTTCTTCTTCC	1959
Db	2273	AAATAGCTTTAGTTT - AAAAAAATCCTCCCTTCCCTTCCCTTCTTCTTCTTCC	2330
QY	1960	TTTTTCCCTTTCCTTTTCCCTTTCCTTTCCCTTCTTCCCTTCCCTTCTTCTT	2019
Db	2331	TTTTTCCCTTTCCTTTTCCCTTTCCTTTCCCTTCTTCCCTTCTTCTTCTT	2390
QY	2020	TTTCTTCCCTTCTGCTGCTGAACTTTTAAAGAGTCTCTAAGTGAAGAGATGGAAGCC	2079
Db	2391	TTTCTTCCCTTCTGCTGCTGAACTTTTAAAGAGTCTCTAAGTGAAGAGATGGAAGCC	2450
QY	2080	AGCCCTGCCAAAGATGGAATCCATAATATGATGCCAGTGAACCTTATTTGAAACATA	2139
Db	2451	AGCCCTGCCAAAGATGGAATCCATAATATGATGCCAGTGAACCTTATTTGAAACATA	2510
QY	2140	CCGTCCCAATGACTAAAGAAATCAAAGAGAGAAACCAAGTTCCCTTAAAGTACAGTGCA	2199
Db	2511	CTGTCCCAATGACTAAAGAAATCAAAGAGAGAAACCAAGTTCCCTTAAAGTACAGTGCA	2570
QY	2200	ACATATACAAATTTGACTGAGTGCAATATTAATTTCAATGGAGCAGCCTCTAATTAGACA	2259
Db	2571	ACATATACAAATTTGACTGAGTGCAATATTAATTTCAATGGAGCAGCCTCTAATTAGACA	2630
QY	2260	ACTTAAAGCAACGTTGCATCGGCTGCTTCTTAATCATTTGCTTTTCCATCTAATCAGTTACA	2319
Db	2631	ACTTAAAGCAACGTTGCATCGGCTGCTTCTTAATCATTTGCTTTTCCATCTAATCAGTTACA	2690
QY	2320	GCCATTTGATTCCTTAAATTTGTTTTTCAAGTCTTCCAGGTATTTGTTAGTTAGCTACTA	2379
Db	2691	GCCATTTGATTCCTTAAATTTGTTTTTCAAGTCTTCCAGGTATTTGTTAGTTAGCTACTA	2750
QY	2380	TGTAACTTTTTCAGGGAATAGTTTAAAGCTTTATTCATTCATGCAATACTTAAAGAAATA	2439
Db	2751	TGTAACTTTTTCAGGGAATAGTTTAAAGCTTTATTCATTCATGCAATACTTAAAGAAATA	2810
QY	2440	AGAACTGCAATTTTGTGCTGCTTGAACAATTAAGAAATTAAGAGCAAAATGA	2499
Db	2811	AGAACTGCAATTTTGTGCTGCTTGAACAATTAAGAAATTAAGAGCAAAATGA	2870
QY	2500	ATCCTGAAGGAAGATTTTAAAAATGTTTTGTTTCTTCTTACAATGGAGATTTTTTGT	2559
Db	2871	ATCCTGAAGGAAGATTTTAAAAATGTTTTGTTTCTTCTTACAATGGAGATTTTTTGT	2930
QY	2560	ACCAGCTTACCACTTTTCAGGCCATTTATTAATATGGGAATTTAACTTACTCAAGCAATA	2619
Db	2931	ACCAGCTTACCACTTTTCAGGCCATTTATTAATATGGGAATTTAACTTACTCAAGCAATA	2990
QY	2620	GTTGAAGGGAAGGTGCATTAATTAATCAACGATGCAATTTATGTTGTGCGCAGTCTGCTCC	2679
Db	2991	GTTGAAGGGAAGGTGCATTAATTAATCAACGATGCAATTTATGTTGTGCGCAGTCTGCTCC	3050
QY	2680	AAACATCAATTTCTTAAATGAGCTCCAGTTTACCTTAAATGTTCACTGACACAAAGATG	2739
Db	3051	AAACATCAATTTCTTAAATGAGCTCCAGTTTACCTTAAATGTTCACTGACACAAAGATG	3110
QY	2740	AGATTACACCTACAGTGACTCTGAGTAGTCAACATATATTAAGCACTGCACATGAGATATAG	2799
Db	3111	AGATTACACCTACAGTGACTCTGAGTAGTCAACATATATTAAGCACTGCACATGAGATATAG	3170
QY	2800	ATCCCTAGAATTTGTCAAGAGTGCACTCTCTACTTTGGGAGGTACAATTTGCCATATGATTT	2859
Db	3171	ATCCCTAGAATTTGTCAAGAGTGCACTCTCTACTTTGGGAGGTACAATTTGCCATATGATTT	3230
QY	2860	CTAGCTGCCATGTGTGTAGGAATGTGATACTGCTGTTTGGCAAAAGTCAACAGACCTTGCC	2919

Db	3231	CTAGCTGCCATGCTGGTTAGGAATGTGATFCTGCTGTTTGCAAAAGTCACAGACCTTGCC	3290
QY	2920	TCAGAAGAGCTGTGAGCCAGTATTCATTTAAGAG	2954
Db	3291	TCAGAAGAGCTGTGAGCCAGTATTCATTTAAGAG	3325
RESULT 9			
AAV47645			
ID	AAV47645	standard; cDNA; 1615 BP.	
XX			
AC	AAV47645;		
XX			
DT	07-DEC-1998	(first entry)	
XX			
DE	Steroid hormone receptor homologue HE8AB36	CDNA.	
KW	HE8AN36; steroid hormone receptor; human; transcription factor;		
KW	inflammation; arthritis; autoimmune disease; diabetes;		
KW	transplant rejection; graft versus host disease; cancer;		
KW	reproductive disorder; obesity; atherosclerosis; gyrate atrophy;		
KW	diagnosis; therapy; ss.		
XX			
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	227..1534	
FT		/*tag= a	
XX			
PN	EP866127-A2.		
XX			
PD	23-SEP-1998.		
XX			
XX	13-MAR-1998;	98EP-00301914.	
XX			
PR	17-MAR-1997;	97GB-00005451.	
PR	16-FEB-1998;	98GB-00003289.	
XX			
PA	(SMIK ) SMITHKLINE BEECHAM PLC.		
XX			
PI	Mathias SL;		
XX			
DR	WPI; 1998-482963/42.		
XX	P-PSDB; AAW61192.		
PT			
PT	New isolated polypeptide(s) are steroid hormone receptor homologues -		
PT	used for treating inflammation, arthritis, auto-immune disease, diabetes,		
PT	transplant rejection etc.		
XX			
PS	Claim 9; Page 12; 20pp; English.		
XX			
CC	This polynucleotide codes for HE8AB36 (see AAW61192), a member of the		
CC	nuclear hormone receptor family of polypeptides that are ligand regulated		
CC	transcription factors involved in the regulation of cellular homeostasis		
CC	and differentiation via the modulation of gene transcription.		
CC	Polynucleotides of the invention can be obtained from a cDNA library		
CC	derived from mRNA of cells of human 8-wk whole embryo and retina, using		
CC	expressed sequence tag analysis. Expression systems, host cells and a		
CC	process for producing HE8AN36 polypeptides are provided. HE8AN36		
CC	polypeptides and polynucleotides may be used in the treatment of chronic		
CC	and acute inflammation, arthritis, autoimmune diseases, diabetes,		
CC	transplant rejection, graft versus host disease, reproductive disorders,		
CC	cancer, obesity, atherosclerosis, gyrate atrophy and other visual		
CC	disorders. The invention also relates to methods for identifying agonists		
CC	and antagonists/inhibitors, and for treating conditions associated with		
CC	HE8AN36 imbalance using such compounds. Diagnostic assays for detecting		
CC	diseases associated with inappropriate HE8AN36 activity or levels are		
CC	also provided		
XX			
XX			
Sequence	1615 BP; 437 A; 395 C; 396 G; 387 T; 0 U; 0 Other;		
Query Match	52.2%; Score 1559; DB 2; Length 1615;		

Best Local Similarity 99.7%; Pred. No. 0;			
Matches 1562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;			
QY	17	GTGGAATCGGCTGTGCTCACTAGAGAAACATTTGTGTTAATTGCACTGTGCTCTGCAAG	76
Db	49	GTGCCATAGGCTGTGCTCACTAGAGAAACATTTGTGTTAATTGCACTGTGCTCTGCAAG	108
QY	77	AAACTTGTATTATAGCTGGGGTGACAAATAATGTTGCCGGTGCACATGATTCCGT	136
Db	109	AAACTTGTATTATAGCTGGGGTGACAAATAATGTTGCCGGTGCACATGATTCCGT	168
QY	137	AGAACTTGGCTTCTGTAATCTTTTCCCTGCACTACGAGAAAGAGCTTCTGCAAGAT	196
Db	169	AGAACTTGGCTTCTGTAATCTTTTCCCTGCACTACGAGAAAGAGCTTCTGCAAGAT	228
QY	197	GTCAACAAGATCGACACATTTGATTCAGCTGTCTCTTCATCAAGACGGAACCTTC	256
Db	229	GTCAACAAGATCGACACATTTGATTCAGCTGTCTCTTCATCAAGACGGAACCTTC	288
QY	257	CAGCCAGCCTCCCTGACGACGCGTCAACCAACACAGCCCTGTGGCTCTTCAGACGC	316
Db	289	CAGCCAGCCTCCCTGACGACGCGTCAACCAACACAGCCCTGTGGCTCTTCAGACGC	348
QY	317	CAGTGGAGCTACAGTTCAACCATGATGGCCATCAGAACGCACTTGACTGCCACCTCT	376
Db	349	CAGTGGAGCTACAGTTCAACCATGATGGCCATCAGAACGCACTTGACTGCCACCTCT	408
QY	377	CTACCCTTCTGCTCCTATCCTTGGAGGTAGTGGCCTGTCAAGAACTGTATGATGACTG	436
Db	409	CTACCCTTCTGCTCCTATCCTTGGAGGTAGTGGCCTGTCAAGAACTGTATGATGACTG	468
QY	437	CTCCAGCACCATTTGTTGAAGATCCCAGACCAAGTGTGAATACATGCTCACTCGATGCC	496
Db	469	CTCCAGCACCATTTGTTGAAGATCCCAGACCAAGTGTGAATACATGCTCACTCGATGCC	528
QY	497	CAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGTACCACTAAGGGGTAGCATC	556
Db	529	CAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGTACCACTAAGGGGTAGCATC	588
QY	557	ATGTGAAGCCTGCAAGGCATTCTTCAAGAGGACAATTCAAGGCAATATAGATACAGCTG	616
Db	589	ATGTGAAGCCTGCAAGGCATTCTTCAAGAGGACAATTCAAGGCAATATAGATACAGCTG	648
QY	617	CCCTGCCACGAATGATGTGAATCACAAAGCCGACAGCTAATCCTGCCAGGCTTGCCG	676
Db	649	CCCTGCCACGAATGATGTGAATCACAAAGCCGACAGCTAATCCTGCCAGGCTTGCCG	708
QY	677	CTTCATGAAGTGTTTAAAGTGGGCATGCTGAAAGAAAGGGGTGCGTCTTGACAGAGTACG	736
Db	709	CTTCATGAAGTGTTTAAAGTGGGCATGCTGAAAGAAAGGGGTGCGTCTTGACAGAGTACG	768
QY	737	TGAGGTCGGCAGAAAGTACAAAGCCGACGATAGATCGGAGAACAGCCCATACCTGAACCC	796
Db	769	TGAGGTCGGCAGAAAGTACAAAGCCGACGATAGATCGGAGAACAGCCCATACCTGAACCC	828
QY	797	TCAGCTGGTTCAGCCAGCCAAAAAGCCATATAACAAGATTGCTCACAATTGTTGGTGGC	856
Db	829	TCAGCTGGTTCAGCCAGCCAAAAAGCCATATAACAAGATTGCTCACAATTGTTGGTGGC	888
QY	857	TGAACCGGAGAGATCTATGCCATGCTGACCTTACTGTCCCGACAGATGACATCAAAAGC	916
Db	889	TGAACCGGAGAGATCTATGCCATGCTGACCTTACTGTCCCGACAGATGACATCAAAAGC	948
QY	917	CCTCACTACACTGTGTACTTGGCCGACCGAGAGTGTGTGTTATCATTTGGATGGCGAA	976
Db	949	CCTCACTACACTGTGTACTTGGCCGACCGAGAGTGTGTGTTATCATTTGGATGGCGAA	1008
QY	977	GCATATTCAGGCTTCTCCACGCTGTCCCTGGCGGACCCAGATGAGCTTCTGCAGAGTGC	1036
Db	1009	GCATATTCAGGCTTCTCCACGCTGTCCCTGGCGGACCCAGATGAGCTTCTGCAGAGTGC	1068
QY	1037	TTGGATGGAATTTTGATCCTTGGTGTGCTATACCGGTCTCTTTTCATTTGAGGATGAAC	1096

Db	1069	TTGGATGGAATTTTGATCCTTGGTGTGCTATACCGGTCTCTTTTCATTTGAGGATGAAC	1128
QY	1097	TGCTATGACAGACGATTATATATATAGACGGAAGCCAGTCCAAATTAGCAGGCTTCTGA	1156
Db	1129	TGCTATGACAGACGATTATATATATAGACGGAAGCCAGTCCAAATTAGCAGGCTTCTGA	1188
QY	1157	TCTAATAATATGCTATCCTCGACGCTGTTAAAGAAATACAAGACATGAAGCTGGAAAAAG	1216
Db	1189	TCTAATAATATGCTATCCTCGACGCTGTTAAAGAAATACAAGACATGAAGCTGGAAAAAG	1248
QY	1217	AGAAATTGTCACCCCTCAAAAGCTATAGCTCTTGCTAATTCAAGACTCCATGCACATAGAAG	1276
Db	1249	AGAAATTGTCACCCCTCAAAAGCTATAGCTCTTGCTAATTCAAGACTCCATGCACATAGAAG	1308
QY	1277	TGTTGAAGCCGTTCAAGAGCTTCAGGATGCTTACATGAAGCGCTGCAGGATTATGAAGC	1336
Db	1309	TGTTGAAGCCGTTCAAGAGCTTCAGGATGCTTACATGAAGCGCTGCAGGATTATGAAGC	1368
QY	1337	TGGCCAGCACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGATGACACTGCCACTCCT	1396
Db	1369	TGGCCAGCACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGATGACACTGCCACTCCT	1428
QY	1397	GAGGACAGCCTCTACCAAGCCGTGACAGATTCTACAACATCAAACTAGAAAGCAAAAGT	1456
Db	1429	GAGGACAGCCTCTACCAAGCCGTGACAGATTCTACAACATCAAACTAGAAAGCAAAAGT	1488
QY	1457	CCCAATGCACAACTTTTTTTGGAATGTTGAGGCCCAAGGTCTGACTTAAAGCTCCCTG	1516
Db	1489	CCCAATGCACAACTTTTTTTGGAATGTTGAGGCCCAAGGTCTGACTTAAAGCTCCCTG	1548
QY	1517	GGCCTTCCCATCTTCATGTTGAAAAAGGAAAAATAAACCCAGAGTGTATGCGAAGAAA	1576
Db	1549	GGCCTTCCCATCTTCATGTTGAAAAAGGAAAAATAAACCCAGAGTGTATGCGAAGAAA	1608
QY	1577	CTTAGAG 1583	
Db	1609	CTTAGAG 1615	
RESULT 10			
AAZ95741			
ID	AAZ95741 standard; cDNA; 1431 BP.		
XX			
AC	AAZ95741;		
XX			
DT	14-JUN-2000 (first entry)		
XX			
DE	Human oestrogen related receptor 3 encoding cDNA SEQ ID NO:2.		
XX			
KW	Human; ligand-combining; oestrogen related receptor; ERR3; diagnosis;		
KW	inflammation; cancer; osteoporosis; diabetes; renal disease; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	JP2000041681-A.		
XX			
PD	15-FEB-2000.		
XX			
PF	31-JUL-1998; 98JP-00217933.		
XX			
PR	31-JUL-1998; 98JP-00217933.		
XX			
PA	(KYOWA) KYOWA HAKKO KOGYO KK.		
XX			
DR	WPI; 2000-295782/26.		
DR	P-PSDB; AAY82263.		
XX			
PT	A new protein comprising a 435 amino acid sequence.		
XX			
PS	Claim 6; Page 19-20; 22pp; Japanese.		
XX			
CC	The present sequence encodes a human oestrogen related receptor 3 (ERR3) protein which has ligand-combining activity. The polynucleotide sequence		

CC encoding ERR3, and fragments of it, can be used for the diagnosis and the  
CC treatment of inflammations, cancers, osteoporosis, diabetes or renal  
CC diseases  
XX  
SQ Sequence 1431 BP; 401 A; 351 C; 348 G; 331 T; 0 U; 0 Other;  
Query Match 47.5%; Score 1418.2; DB 3; Length 1431;  
Best Local Similarity 99.8%; Pred. No. 1.1e-312;  
Matches 1420; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 176 GGAAGAGCTTCTCTGAGAAATGTCAAACAAGATCGACACATGATTCAGCTGTTGTC 235  
DB 9 GGACAGGCTTCTCTGAGAAATGTCAAACAAGATCGACACATGATTCAGCTGTTGTC 68  
QY 236 CTTCAATCAAGACGAACCTTCCAGCCCGCTCCCTGACGACAGCGTCAACCAACAG 295  
DB 69 CTTCAATCAAGACGAACCTTCCAGCCCGCTCCCTGACGACAGCGTCAACCAACAG 128  
QY 296 CCCTGTGGCTTTCAGACGCCAGTGGAGCTACAGTTCACCATGATGGCCATCAGAA 355  
DB 129 CCCTGTGGCTTTCAGACGCCAGTGGAGCTACAGTTCACCATGATGGCCATCAGAA 188  
QY 356 CGGACTTGACTCGCCACCTCTCTACCCCTTCTCTCTATCCTGGAGAGTATGGCCCTGT 415  
DB 189 CGGACTTGACTCGCCACCTCTCTACCCCTTCTCTCTATCCTGGAGAGTATGGCCCTGT 248  
QY 416 CAGAAACTGTATGATGACTGCTCCAGCACCAATGTTGAAGATCCCCAGACCAAGTGTGA 475  
DB 249 CAGAAACTGTATGATGACTGCTCCAGCACCAATGTTGAAGATCCCCAGACCAAGTGTGA 308  
QY 476 ATACATGCTCAACTCGATGCCCAAGAGACTGTGTTAGTGTGTGATGACATCGCTTCTGG 535  
DB 309 ATACATGCTCAACTCGATGCCCAAGAGACTGTGTTAGTGTGTGATGACATCGCTTCTGG 368  
QY 536 GTACCACTATGGGTAGCATCATGTGAAGCCTGCAAGCATTTCTCAAGAGACAATTCA 595  
DB 369 GTACCACTATGGGTAGCATCATGTGAAGCCTGCAAGCATTTCTCAAGAGACAATTCA 428  
QY 596 AGGCAATATAGAAATACAGCTGCCCTGCCACGAATGATGAAATCACAAAGCGCAGACG 655  
DB 429 AGGCAATATAGAAATACAGCTGCCCTGCCACGAATGATGAAATCACAAAGCGCAGACG 488  
QY 656 TAAATCCTGCCAGGCTTGGCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGG 715  
DB 489 TAAATCCTGCCAGGCTTGGCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGG 548  
QY 716 GGTGCGTCTTGACAGAGTACGTGAGGTGCGGACGAAGTACAAGCGCAGATAGATCGGGA 775  
DB 549 GGTGCGTCTTGACAGAGTACGTGAGGTGCGGACGAAGTACAAGCGCAGATAGATCGGGA 608  
QY 776 GAACAGCCCATACCTGAACCTCAGCTGTTCAAGCCCAAAAAAGCCATATTAACAAGAT 835  
DB 609 GAACAGCCCATACCTGAACCTCAGCTGTTCAAGCCCAAAAAAGCCATATTAACAAGAT 668  
QY 836 TGTCTCACATTTGTTGGTGTGAACCGGAGAGATCTATGCCATGCTGACCCCTACTGT 895  
DB 669 TGTCTCACATTTGTTGGTGTGAACCGGAGAGATCTATGCCATGCTGACCCCTACTGT 728  
QY 896 CCCCAGACAGTACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGT 955  
DB 729 CCCCAGACAGTACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGT 788  
QY 956 GGTATCAATTGGATGGGCGAAGCATATTCAGGCTTCTCAAGCTGTCCCTGGCGGACCA 1015  
DB 789 GGTATCAATTGGATGGGCGAAGCATATTCAGGCTTCTCAAGCTGTCCCTGGCGGACCA 848  
QY 1016 GATGAGCCTTCTGACAGAGTGTGATGAAATTTTGAATCCTTGTGCTATACCGGTG 1075  
DB 849 GATGAGCCTTCTGACAGAGTGTGATGAAATTTTGAATCCTTGTGCTATACCGGTG 908  
QY 1076 TCTTCAATTGAGGATGAACCTTGTCTATGACAGAGATTAATAATGACGAAGACCAAGTC 1135  
DB 909 TCTTCAATTGAGGATGAACCTTGTCTATGACAGAGATTAATAATGACGAAGACCAAGTC 968

QY 1136 CAATTTAGCAGGCTTCTTGAATCTAAATATGCTATCCTGACGCTGTAAGAATAACAA 1195  
DB 969 CAATTTAGCAGGCTTCTTGAATCTAAATATGCTATCCTGACGCTGTAAGAATAACAA 1028  
QY 1196 GAGCATGAAGCTGGAAGAAAGAAATTTGTCAACCCCTCAAGCTATAGCTCTTGCTAATTG 1255  
DB 1029 GAGCATGAAGCTGGAAGAAAGAAATTTGTCAACCCCTCAAGCTATAGCTCTTGCTAATTG 1088  
QY 1256 AGACTCCATGACATAGAAGATGTGAAGCCGTTCAAGAGCTTCAAGATGCTTACATGA 1315  
DB 1089 AGACTCCATGACATAGAAGATGTGAAGCCGTTCAAGAGCTTCAAGATGCTTACATGA 1148  
QY 1316 AGCGCTGCAGATTTATGAAGCTGGCCAGACATGGAAGACCTGCTGAGAGCTGCAAGAT 1375  
DB 1149 AGCGCTGCAGATTTATGAAGCTGGCCAGACATGGAAGACCTGCTGAGAGCTGCAAGAT 1208  
QY 1376 GCTGATGACACTGCCACTCCTGAGGCGACACCTTACCAAGGCGGTGACATTTCTACAA 1435  
DB 1209 GCTGATGACACTGCCACTCCTGAGGCGACACCTTACCAAGGCGGTGACATTTCTACAA 1268  
QY 1436 CATCAAACTAGAAAGGCAAAAGTCCCAATGCACAAACTTTTGTGAAATGTGAGGCCAA 1495  
DB 1269 CATCAAACTAGAAAGGCAAAAGTCCCAATGCACAAACTTTTGTGAAATGTGAGGCCAA 1328  
QY 1496 GGTCTGACTAAAGCTCCCTGGGCTTCCATCCTCATGTTGAAGAAAGGAAATTAAC 1555  
DB 1329 GGTCTGACTAAAGCTCCCTGGGCTTCCATCCTCATGTTGAAGAAAGGAAATTAAC 1388  
QY 1556 CCAAGAGTGAATGCAAGAAACTTAGAGTTAGTTAACAACAT 1598  
DB 1389 CCAAGAGTGAATGCAAGAAACTTAGAGTTAGTTAACAACAT 1431  
RESULT 11  
AD23488  
ID AD23488 standard; DNA; 1377 BP.  
XX  
AC AD23488;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE DNA encoding the ligand binding domain of the ERR3 protein.  
XX  
XX  
KW oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;  
KW ligand binding domain; fertility; birth control; bone remodeling;  
KW cancer; protein coordinate data; gene; ds.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1377  
FT /tag= a  
FT /product= "Ligand binding domain of the ERR3 protein"  
XX  
XX WO2003064468-A2.  
XX  
XX 07-AUG-2003.  
XX  
XX 30-JAN-2003; 2003WO-EP000959.  
XX  
XX 31-JAN-2002; 2002US-0352551P.  
XX  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX Moras D, Renaud J, Greschik H, Wurtz J;  
XX  
XX WPI; 2003-663467/62.  
XX  
XX P-PSDB; AD23486.  
XX  
XX New peptide fragment, useful for screening compounds that are agonists or  
XX antagonists of the transcriptional-activating activity of the estrogen-  
XX related receptor 3 (ERR3).



XX PS Disclosure; SEQ ID NO 3; 259pp; English.  
XX CC This invention relates to novel peptide fragments that have an agonistic  
CC or antagonistic effect on the transcriptional-activating activity of the  
CC oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen  
CC related receptor gamma protein. Specifically, the peptide fragment of the  
CC invention comprises a ligand binding domain of ERR3, which becomes  
CC functionally active when fused to a protein containing a DNA binding  
CC domain and mimics the transcriptional-activating activity of the complete  
CC ERR3 protein. As such, the peptide fragment can be used in a screening  
CC method or to design and select compounds that affect ERR3 activity.  
CC Furthermore, the agonists and antagonists of ERR3 are biologically active  
CC compounds that can be used to modulate the oestrogenic response on  
CC fertility, birth control, bone remodelling, breast and prostate cancer.  
CC This polynucleotide sequence is the DNA encoding the ligand binding  
CC pocket of the ERR3 protein of the invention. NOTE: This sequence is given  
CC in the sequence listing but is not further referred to in the  
CC specification.  
XX CC  
SQ Sequence 1377 BP; 375 A; 341 C; 335 G; 326 T; 0 U; 0 Other;

Query Match 46.1%; Score 1377; DB 9; Length 1377;  
Best Local Similarity 100.0%; Pred. No. 2.6e-303;  
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 ATGATTCGGTAGAATCTTCCCTTCTGATCTTTCCCTGCACTAGAGAGAGCTT 185  
Db 1 ATGATTCGGTAGAATCTTCCCTTCTGATCTTTCCCTGCACTAGAGAGAGCTT 60

QY 186 CTCTGAGAAATGTCAACAAGATCGACACATGATTCAGCTGTTCCTTCATCAAG 245  
Db 61 CTCTGAGAAATGTCAACAAGATCGACACATGATTCAGCTGTTCCTTCATCAAG 120

QY 246 ACGGAACCTTCCAGCCCGACCTCCCTGACGAGACGGTCAACCAACCAAGCCCTGTGGC 305  
Db 121 ACGGAACCTTCCAGCCCGACCTCCCTGACGAGACGGTCAACCAACCAAGCCCTGTGGC 180

QY 306 TCTTCAGACGCCAGTGGAGAGCTACAGTTCACCAATGAATGGCCATCAGAACGCACTTGAC 365  
Db 181 TCTTCAGACGCCAGTGGAGAGCTACAGTTCACCAATGAATGGCCATCAGAACGCACTTGAC 240

QY 366 TCGCCACCTCTTACCTTCTGCTCTCTATCCTGGAGGTAGTGGCCCTGCAGAAAATG 425  
Db 241 TCGCCACCTCTTACCTTCTGCTCTCTATCCTGGAGGTAGTGGCCCTGCAGAAAATG 300

QY 426 TATGATGACTGCTCCAGCACCATTTGTTGAAGATCCCAAGCAAGTGAATATACATGCTC 485  
Db 301 TATGATGACTGCTCCAGCACCATTTGTTGAAGATCCCAAGCAAGTGAATATACATGCTC 360

QY 486 AACTCGATGCCCAAGAGACTGTGTTAGTGTGTGATCGCTTCTGGTACCACCTAT 545  
Db 361 AACTCGATGCCCAAGAGACTGTGTTAGTGTGTGATCGCTTCTGGTACCACCTAT 420

QY 546 GGGGTAGCATCATGTGAAGCCTGCAAGGCAATTCTCAAGAGACAATTAAGGCAATATA 605  
Db 421 GGGGTAGCATCATGTGAAGCCTGCAAGGCAATTCTCAAGAGACAATTAAGGCAATATA 480

QY 606 GAATACAGCTGCCCTGCCACGAATGAATGTAAATCACAAAAGCGCAGAGTAAATCCTGC 665  
Db 481 GAATACAGCTGCCCTGCCACGAATGAATGTAAATCACAAAAGCGCAGAGTAAATCCTGC 540

QY 666 CAGGCTTGGCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGGGTGCGTCTT 725  
Db 541 CAGGCTTGGCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGGGTGCGTCTT 600

QY 726 GACAGAGTACGTGAGGTGCGCAGAAATCAAGCGCAGATAGATGCGGAGAAACAGCCCA 785  
Db 601 GACAGAGTACGTGAGGTGCGCAGAAATCAAGCGCAGATAGATGCGGAGAAACAGCCCA 660

QY 786 TACCTGAACCTCAGCTGTTGAGCCAGCAAAAAGCCATATATACAGATTTGTTACAT 845  
Db 661 TACCTGAACCTCAGCTGTTGAGCCAGCAAAAAGCCATATATACAGATTTGTTACAT 720

QY 846 TTGTTGTGGCTGAACCGGAGAAAGATCTATGCCATGCTGACCCCTACTGTCCCCGACAGT 905  
Db 721 TTGTTGTGGCTGAACCGGAGAAAGATCTATGCCATGCTGACCCCTACTGTCCCCGACAGT 780

QY 906 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACGAGAGTTGGTGTATCATTT 965  
Db 781 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACGAGAGTTGGTGTATCATTT 840

QY 966 GATGGGGGGAAGCATATATCCAGGCTTCTCCACGCTGTCCCTGGCGGACCAAGATGAGCCTT 1025  
Db 841 GATGGGGGGAAGCATATATCCAGGCTTCTCCACGCTGTCCCTGGCGGACCAAGATGAGCCTT 900

QY 1026 CTCAGAGTGTGATGAAATTTGATCCTTGGTGTCTATACCGGTCTCTTCAATTT 1085  
Db 901 CTCAGAGTGTGATGAAATTTGATCCTTGGTGTCTATACCGGTCTCTTCAATTT 960

QY 1086 GAGGATGAACCTGTCTATGACAGACGATTATATATGACGAAAGACCAAGTCCAAATTAACA 1145  
Db 961 GAGGATGAACCTGTCTATGACAGACGATTATATATGACGAAAGACCAAGTCCAAATTAACA 1020

QY 1146 GGCCCTTCTGATCTAAATTAATGCTATCCTGACGCTGTGTAAGAAATACAGAGCATGAAG 1205  
Db 1021 GGCCCTTCTGATCTAAATTAATGCTATCCTGACGCTGTGTAAGAAATACAGAGCATGAAG 1080

QY 1206 CTGAAAAAAGAAATTTGTCAACCTCAAGCTAATAGCTCTTGCTAATTCAGACTCCATG 1265  
Db 1081 CTGAAAAAAGAAATTTGTCAACCTCAAGCTAATAGCTCTTGCTAATTCAGACTCCATG 1140

QY 1266 CACATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAGAGATGCTTACATGAAGCGTCCAG 1325  
Db 1141 CACATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAGAGATGCTTACATGAAGCGTCCAG 1200

QY 1326 GATTATGAAGCTGGCCAGCACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGATGACA 1385  
Db 1201 GATTATGAAGCTGGCCAGCACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGATGACA 1260

QY 1386 CTGCCACTCCTGAGGCAAGACCTCTACCAAGGCGGTGACGATTTCTACAAATCAAACTA 1445  
Db 1261 CTGCCACTCCTGAGGCAAGACCTCTACCAAGGCGGTGACGATTTCTACAAATCAAACTA 1320

QY 1446 GAAGGCAAAAGTCCCAATGCACAAACTTTTGGAAATGTTGGAGGCCAAGGTCTGA 1502  
Db 1321 GAAGGCAAAAGTCCCAATGCACAAACTTTTGGAAATGTTGGAGGCCAAGGTCTGA 1377

RESULT 12  
AAA40078  
ID AAA40078 standard; cDNA; 1377 BP.  
XX AC AAA40078;  
XX AC  
XX DT 19-OCT-2000 (first entry)  
XX DE Human ERRgamma coding region cDNA.  
XX KW Human; ERRgamma; brain; estrogen-related receptor gamma; cytostatic;  
KW lipid metabolism; fatty acid synthesis; antiarteriosclerotic; treatment;  
KW drug development; diabetes; ss.  
XX OS Homo sapiens.  
XX PN WO200026365-A1.  
XX PD 11-MAY-2000.  
XX PF 02-NOV-1999; 99WO-JP006097.  
XX PR 04-NOV-1998; 98JP-00313194.  
XX PA (KAZU-) KAZUSA DNA RES INST FOUND.  
XX PA (TAIS ) TAISHO PHARM CO LTD.  
XX XX

PI Ohara O, Nagase T, Nomura N, Takayama K, Toyoda H, Yoshimoto M;  
XX WPI; 2000-365614/31.  
DR P-PSDB; AAB09965.  
XX  
PT An estrogen-related receptor gamma protein with lipid metabolism  
PT regulatory and fatty acid synthesis functions, and its encoding gene,  
PT useful as drugs and in developing drugs for treatment of e.g.  
PT arteriosclerosis and diabetes.  
XX  
PS Claim 2a; Page 24-26; 38pp; Japanese.  
XX  
CC This invention describes a novel estrogen-related receptor gamma  
CC (ERGamma) protein (I) and its variants maintaining receptor activity,  
CC having lipid metabolism regulatory and fatty acid synthesis functions.  
CC The product of the invention has cytosolic and antiarteriosclerotic  
CC activity. The protein and gene are useful as drugs and in developing  
CC drugs for treatment of e.g. arteriosclerosis and diabetes. The protein is  
CC obtained by cloning human brain-originated cDNA library. This sequence  
CC encodes the human brain ERGamma protein which is described in the method  
CC of the invention  
XX  
SQ Sequence 1377 BP; 374 A; 341 C; 336 G; 326 T; 0 U; 0 Other;  
Query Match 46.1%; Score 1375.4; DB 3; Length 1377;  
Best Local Similarity 99.9%; Pred. No. 5.9e-303;  
Matches 1376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 126 ATGGATTCGGTAGAAGCTTTGCTTCCCTGAAATCTTTTCCCTGCACTACGAGAAAGAGCTT 185  
Db 1 ATGGATTCGGTAGAAGCTTTGCTTCCCTGAAATCTTTTCCCTGCACTACGAGAAAGAGCTT 60  
QY 186 CTCTGCAGAATGTCAAAACAAGATGCACACATGATCCAGCTGTTCCTTCATCAAG 245  
Db 61 CTCTGCAGAATGTCAAAACAAGATGCACACATGATCCAGCTGTTCCTTCATCAAG 120  
QY 246 ACGGAACCTTCCAGCCCAAGCCTCCCTGACGAGCAGCGTCAACCAACCAAGCCCTGGTGGC 305  
Db 121 ACGGAACCTTCCAGCCCAAGCCTCCCTGACGAGCAGCGTCAACCAACCAAGCCCTGGTGGC 180  
QY 306 TCTTCAGAGCGCAGTGGAGCTACAGTCAACCATGAATGGCCATCAGAACGGAAGCTTGAC 365  
Db 181 TCTTCAGAGCGCAGTGGAGCTACAGTCAACCATGAATGGCCATCAGAACGGAAGCTTGAC 240  
QY 366 TCGCCACCTCTCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 425  
Db 241 TCGCCACCTCTCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
QY 426 TATGATGACTGCTCCAGACCATTTGTTGAAGATCCCAAGTGTGAATACATGCTC 485  
Db 301 TATGATGACTGCTCCAGACCATTTGTTGAAGATCCCAAGTGTGAATACATGCTC 360  
QY 486 AACTCGATGCCCAAGACTGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 545  
Db 361 AACTCGATGCCCAAGACTGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
QY 546 GGGGTAGCATCATGTGAAGCCTGCAAGGCATTTCTCAAGAGCAATTCAGAGCAATATA 605  
Db 421 GGGGTAGCATCATGTGAAGCCTGCAAGGCATTTCTCAAGAGCAATTCAGAGCAATATA 480  
QY 606 GAATACAGCTGCCCTGCCAGCAATGAATGTGAATACAAAGCGCAGACGTAATCTCTGC 665  
Db 481 GAATACAGCTGCCCTGCCAGCAATGAATGTGAATACAAAGCGCAGACGTAATCTCTGC 540  
QY 666 CAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGGGTGCTCTT 725  
Db 541 CAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGGGTGCTCTT 600  
QY 726 GACAGAGTACGTGAGGTGGGAGAGTACAAAGCGCAGAGTAGATGCGAGAACAGCCCA 785  
Db 601 GACAGAGTACGTGAGGTGGGAGAGTACAAAGCGCAGAGTAGATGCGAGAACAGCCCA 660  
QY 786 TACCTGAACCTCAGCTGGTTCAGCCAGCCAAAAGCCATATAACAAGATTGTCTCAT 845

Db 661 TACCTGAACCTCAGCTGGTTCAGCCAGCCAAAAGCCATATAACAAGATTGTCTCAT 720  
QY 846 TTGTTGGTGGCTGAACCCGAGAGATCTATGCCATGCTTGACCTACTGTCCCGACAGT 905  
Db 721 TTGTTGGTGGCTGAACCCGAGAGATCTATGCCATGCTTGACCTACTGTCCCGACAGT 780  
QY 906 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGTATCAT 965  
Db 781 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGTATCAT 840  
QY 966 GGATGGGCGAAGCATATATTCAGAGCTTCTCCAGCGTGTCCCTGGCGGACAGATGAGCCTT 1025  
Db 841 GGATGGGCGAAGCATATATTCAGAGCTTCTCCAGCGTGTCCCTGGCGGACAGATGAGCCTT 900  
QY 1026 CTGCAGAGTGTGATGGAATTTGATTCCTTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1085  
Db 901 CTGCAGAGTGTGATGGAATTTGATTCCTTGGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
QY 1086 GAGGATGAAGCTGTCTATGACAGAGATTAATATATGAGCAGAAAGACAGTCCAAATTAGCA 1145  
Db 961 GAGGATGAAGCTGTCTATGACAGAGATTAATATATGAGCAGAAAGACAGTCCAAATTAGCA 1020  
QY 1146 GGCTTCTTGAATCTAATATATGCTATCTGACAGCTGTGTAAGAAATACAGAGCATGAAG 1205  
Db 1021 GGCTTCTTGAATCTAATATATGCTATCTGACAGCTGTGTAAGAAATACAGAGCATGAAG 1080  
QY 1206 CTGAAAAAGAGAAATTTGTCAACCTCAAGCTATAGCTCTTGTCTAATTGAGACTCCATG 1265  
Db 1081 CTGAAAAAGAGAAATTTGTCAACCTCAAGCTATAGCTCTTGTCTAATTGAGACTCCATG 1140  
QY 1266 CACATAGAGATGTTGAAGCCGTTCAAGAGCTTCAAGATGCTTACATGAAGCGCTGCAG 1325  
Db 1141 CACATAGAGATGTTGAAGCCGTTCAAGAGCTTCAAGATGCTTACATGAAGCGCTGCAG 1200  
QY 1326 GATTATGAAGCTGGCCAGCAGCATGGAAGACCCCTGTCGAGCTGGCAAGATGCTGATGACA 1385  
Db 1201 GATTATGAAGCTGGCCAGCAGCATGGAAGACCCCTGTCGAGCTGGCAAGATGCTGATGACA 1260  
QY 1386 CTGCCACTCTGAGGCGACAGCTCTACCAAGGCGTGACAGATTTCTACACATCAAACTA 1445  
Db 1261 CTGCCACTCTGAGGCGACAGCTCTACCAAGGCGTGACAGATTTCTACACATCAAACTA 1320  
QY 1446 GAAGGCAAAAGTCCCAATGCACAAACTTTTGGAAATGTTGAGGCCAAGCTCTGA 1502  
Db 1321 GAAGGCAAAAGTCCCAATGCACAAACTTTTGGAAATGTTGAGGCCAAGCTCTGA 1377  
RESULT 13  
AAA72605  
ID AAA72605 standard; cDNA; 1996 BP.  
XX  
AC AAA72605;  
XX  
DT 27-NOV-2000 (first entry)  
XX  
DE Human oestrogen related receptor 4 (ERR4) cDNA sequence.  
XX  
KW Oestrogen related receptor 4; ERR4; antiinflammatory; cytosolic;  
KW antidiabetic; nephrotropic; osteopathic; inflammation; cancer;  
KW osteoporosis; diabetes; kidney disease; ss.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 140..1441  
FT /\*tag= a  
FT /product= "ERR4"  
FT /note= "Oestrogen related receptor 4"  
XX  
PN WO200042180-A1.  
XX  
PD 20-JUL-2000.

XX 12-JAN-2000; 2000WO-JP000106.  
XX 14-JAN-1999; 99JP-00007226.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Yamada Y, Nakagawa S, Sekine S;  
XX WPI; 2000-491056/43.  
XX P-PSDB; AAB12970.  
PT Novel protein with ability to bind to a ligand and to DNA, useful in the  
PT diagnosis of and screening of candidate drugs for the treatment of  
PT inflammation, cancer, osteoporosis, diabetes and kidney diseases.  
XX Claim 5; Page 46-51; 57pp; Japanese.  
XX This sequence represents cDNA encoding the human oestrogen related  
XX receptor 4 (ERR4). The invention relates to the ERR4 protein and DNA  
XX sequences, a vector containing the DNA sequence, a transformant  
XX containing the vector, and an antibody specific for the ERR4 protein. The  
XX ERR4 protein exhibits antiinflammatory, cytostatic, antidiabetic,  
XX nephrotropic, and osteopathic activity. The protein, DNA and antibody are  
XX useful in the diagnosis of inflammation, cancer, osteoporosis, diabetes  
XX and kidney disease, and also in the screening of drugs for use in the  
XX treatment of these diseases  
SQ Sequence 1996 BP; 401 A; 633 C; 584 G; 378 T; 0 U; 0 Other;

Query Match 24.1%; Score 719.6; DB 3; Length 1996;  
Best Local Similarity 71.6%; Pred. No. 1.4e-153;  
Matches 960; Conservative 0; Mismatches 374; Indels 6; Gaps 1;

QY 163 CCCTGCACTACGAGAGAGCTTCTCTGAGATGTCACAAAGATCGACACATTGATT 222  
Db 108 CCTCGGCTACCAACACAGCTGCTGAACAGAGATGCTCTCGAACGACAGGCACTGGGCT 167  
QY 223 CCAGCTGTTCTCTCTTCACTCAAGACGGAACCTTCCAGCCCAAGCCTCCCTGACGGAAGCG 282  
Db 168 CCAGCTGCGGCTCTTCACTCAAGACTGAGCCCTCCAGCCCTCTCGGCAATGATGCC 227  
QY 283 TCAACCAACACAGCCCTGCTGCTCTTCAAGCCAGTGGAGCTACAGTCAACCATGA 342  
Db 228 TCAGCCACACAGCCCAAGTGGCTGCTCCAGCCCAAGCGGCTTGGCTGGCCCTGG 287  
QY 343 ATGGCCATCAGACGGAAGCTTGAATGCTGACCTCTTACCTTCTGCTCCATCCCTGGAG 402  
Db 288 GCACCCACGCAACGCTTGAAGTCCGACCC-----CATGTTGCAAGCGCGGGGCTGG 341  
QY 403 GTAGTGGGCTGTGAGGAACTGTATGATGACTGCTCCAGCACCATTGTTGAAGATCCCC 462  
Db 342 GAGGCACCCCATGCGGCAAGAGCTACGAGAGCTGTGCCAGCGGCATCATGAGGAAGACTCGG 401  
QY 463 AGACCAAGTGTGAATACATGCTCACTCGATGCCCAAGAGACTGTGTTAGTGTGTG 522  
Db 402 CCATCAAGTGCAGATACATGCTCAACGCCCATCCCAAGCGCCTGTGCTGTGTGCGGG 461  
QY 523 ACATGCTTCTGGGTAACCATATGGGGTAGATCATGTGAAGCCTGCAAGGCATTCTTCA 582  
Db 462 ACATTGCTCTGGCTACCACTACGGCGTGGCTCTGCGAGGCTTGGCAAGGCTTCTTCA 521  
QY 583 AGAGGACAATTCAAGGCAATATAGAAATACAGCTGCCCTGCCACGAATGATGAATCA 642  
Db 522 AGAGACTATCAAGGGAACATGAGTACAGCTGCCCGGCCACCAACGATGCGAGATCA 581  
QY 643 CAAAGCGCAGAGCTAAATCTGCCAGGCTGCCGCTTCATGAAGTGTAAAGTGGCA 702  
Db 582 CCAACGAGGCGCAAGTCTCTGCCAGGCTGCCGCTTCATGAATGCTCAAGTGGGA 641  
QY 703 TGCTGAAGAGAGGGGTGCTTTGACAGAGTACGTGAGGTCCGCAAGTACAAAGCGCA 762  
Db 642 TGCTGAAGAGAGGTGTGCGCCTTGATCGAGTGGCTGGAGGCGGTCAAGAAATACAAAGCGAC 701

QY 763 GGATAGATGCGGAGAACAGCCCATACCTGAAACCCCTCACTGCTGTTCAAGCCAGCCAAAAGC 822  
Db 702 GGCTGACTCAGAGAGAGAGCCCATACCTGAGCTTACAAATTTCTCCACCTGTAAAGC 761  
QY 823 CATATAACAAGATTGCTCACTTTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGC 882  
Db 762 CATTGACCAAGATTGCTCATACCTACTGCTGGCTGAGCCGGAACAAGCTCTATGCCATGC 821  
QY 883 CTGACCCCTACTGTCCCGACAGTGAATCAAAAGCCCTCACTACTGTGTGACTTGGCCG 942  
Db 822 CTCCCCCTGATGCTGAGGGGAGACATCAAGGCCCTGACCACTCTGTGACCTGGCAG 881  
QY 943 ACCGAGAGTGTGTTATCATTTGATGGGCGGAGCATATTCAGGCTTCTCCACGCTGT 1002  
Db 882 ACCGAGAGCTGTGTTATCATTTGCTGGGCGGAGCATATCCAGGCTTCTCAAGCTCT 941  
QY 1003 CCCTGGCGGACCAATGAGCCTTCTGACAGAGTCTTGAAATTTGATCCTTGGTG 1062  
Db 942 CCTTGGGAGACCAATGAGCCTTCTGACAGAGTCTTGAAATTTGATCCTTGGTG 1001  
QY 1063 TCGTATACCGGTCTTTCATTTGAGATGAATGCTGCTATGACAGAGATTAATATAG 1122  
Db 1002 TCGTATACCGGTCTTTCATTTGAGATGAATGCTGCTATGACAGAGATTAATATAG 1061  
QY 1123 ACGAAGACCAATCAAAATTAAGAGCCTTCTTGAATTAATGCTATGCTGACGCTGG 1182  
Db 1062 ATGAGAGACACTCCGCTTCCGCGGCTGCGGCTGCTGAGCTCTACCGGCATCCTGACGCTGG 1121  
QY 1183 TAAAGAAATACAGAGCATGAAGCTGGAAGAAAGAAATTTGTCACCCCTCAAGCTATAG 1242  
Db 1122 TAGCAGATACAGAGAGCTCAAGGTGAGAGAGAGATTGTGACGCTCAAGGCTCG 1181  
QY 1243 CTCTTCTAATTCAGACTCCATGCATAGATGAAGTGTGAAGCCGTTCAGAGCTTCAGG 1302  
Db 1182 CCTCGCAACTCCGATTTCCATGTACATCGAGATCTAGAGGCTGTCCAGAGCTGCAGG 1241  
QY 1303 ATGCTTACATGAAGCGCTGCAGGATTTATGAAGCTGGCCAGCATGGAAGCCCTGCTC 1362  
Db 1242 ACCTGCTGACAGAGCACTGACAGACTACGAGCTGAGCCAGCCATGAGAGCCCTGGA 1301  
QY 1363 GAGCTGGCAAGATGCTGATGACACTGCACTGCTGAGGCGAGACCTTACAGAGCCGTGC 1422  
Db 1302 GGACGGGCAAGCTGCTGCTGACACTGCGCTGCTGCGGAGAGCGCGCCAAAGCCGTGC 1361  
QY 1423 AGCATTTCTACACATCAAACTAGAAAGGCAAGTCCCAATGCACAAACTTTTGGAAA 1482  
Db 1362 AGCACTTCTATAGCGTCAAACTGACAGGCGAAAGTGCATGACAAACTCTTCTGGAGA 1421  
QY 1483 TGTGAGGCGCAAGTCTGA 1502  
Db 1422 TGCTGAGGCGCAAGTGTGA 1441

RESULT 14  
AAX32265  
ID AAX32265 standard; DNA; 2807 BP.  
XX AAX32265;  
AC AAX32265;  
XX 15-JUN-1999 (first entry)  
DT 15-JUN-1999 (first entry)  
XX Human nuclear receptor protein nNr1 encoding DNA.  
DE Human nuclear receptor; nNr1; nNr2; cell differentiation; human; ss.  
XX Homo sapiens.  
OS Homo sapiens.  
XX MO9910367-A1.  
PN MO9910367-A1.  
XX 04-MAR-1999.  
PD 04-MAR-1999.  
XX 27-AUG-1998; 98WO-US017826.  
PF 27-AUG-1998; 98WO-US017826.



XX 27-AUG-1997; 97US-0057090P.  
PR 21-OCT-1997; 97US-0062902P.  
PR 19-MAR-1998; 98US-0078633P.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Chen F;  
XX  
DR WPI: 1999-190586/16.  
DR P-PSDB; AAY03836.  
XX

PT Newly purified DNA polynucleotides encoding human nuclear trans-acting  
PT receptor proteins - useful in the diagnosis, treatment and prophylaxis of  
PT cell differentiation, development and physiological function.

PS Claim 9; Page 41-42; 82pp; English.

XX The invention relates to DNA molecules encoding human nuclear receptor  
CC (nNR) proteins nNR1 and nNR2. The nNR proteins (including mutants and/or  
CC fragments) form pharmaceutical compositions that are useful in the  
CC diagnosis, treatment and prophylaxis of cell differentiation, development  
CC and physiological function. The proteins are also useful for identifying  
CC downstream target genes and ligands regulating their activity. In  
CC particular, fusion constructs (especially glutathione S-transferase (GST)  
CC -nNR1 and/or GST-nNR2) expressing fusion proteins are useful in screening  
CC for (ant)agonists that are useful as modulators in cell differentiation,  
CC development and physiological function. Antibodies for the nNR proteins  
CC are useful for measuring the levels of these proteins. The present  
CC sequence represents a DNA encoding a human nNR1 protein

SQ Sequence 2807 BP; 701 A; 777 C; 770 G; 559 T; 0 U; 0 Other;

Query Match 24.0%; Score 715; DB 2; Length 2807;  
Best Local Similarity 71.5%; Pred. No. 1.8e-152;  
Matches 956; Conservative 0; Mismatches 375; Indels 6; Gaps 1;

QY 179 AGAGCTTCTCTGCAGAATGTCAACAAGATCGACACATTGATTCAGCTGTTCCTT 238  
Db 934 AGGGCTGTGAAACAGATGTCTCGAGCAAGCACCTGGGCTCCAGCTCGGCTCCTT 993  
QY 239 CATCAAGACGGAACCTTCCAGCCCGCCTCCCTGACGACAGCGCTCAACCAACAGAGCC 298  
Db 994 CATCAAGACTGAGCCGTCCAGCCCGTCTCGGGCATAGATGCCCTCAGCCACAGAGCC 1053  
QY 299 TGGTGGCTCTTCAGACGCGCAGTGGAGCTACAGTTCACCATGAATGGCCATCAGAACGG 358  
Db 1054 CAGTGGCTGTCCGACGCGCGGCTTTGGCCTGGCCCTGGGCAACCAAGCCAGCG 1113  
QY 359 ACTTGACTCGCCACCTCTTACCCCTTCTGCTCTATCCTGGAGGTAGTGGGCTGTAG 418  
Db 1114 TCTGACTCGCCACC-----CATGTTTGAGGCGCGGCTGGAGGCAACCCATGCCG 1167  
QY 419 GAAACTGTATGATGACTGCTCCAGCACCATTGTTGAAGATCCCAAGCAAGTGTGAATA 478  
Db 1168 CAAGAGCTACGAGGACTGTGCCAGCGGCATCATGAGGACTCGGCCATCAAGTGCAGTA 1227  
QY 479 CATGCTCAACTCGATGATGAGCGCTGCAAGCATTTCTTCAAGAGCAATTCAGG 538  
Db 1228 CATGCTCAAGCCATCCCAAGCGCTGTGCTCTGTGCGGGGACATTTGCTGTGCTA 1287  
QY 539 CCACTATGGGGTAGCATCATGTGAAGCGCTGCAAGCATTTCTTCAAGAGCAATTCAGG 598  
Db 1288 CCACTACGGCGGTGGCTCTCTGCGAGGCTTGCAAGGCTTCTTCAAGAGCAATTCAGG 1347  
QY 599 CAATATAGAATACAGCTGCCCTGCCAGAAATGAATGTGAATCAACAAGCGACAGCTAA 658  
Db 1348 GAACATGTGATACAGCTGCCCGGCCAACAGAGTGCAGATCACCAACGAGGCGCAA 1407  
QY 659 ATCCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGGGT 718  
Db 1408 GTCTGCGCAGGCTGCGGCTTCATGAATGCTCAAAAGTGGGATGCTGAAGAAGGTGT 1467

QY 719 GCGTCTTGACAGAGTACGTGAGGTCCGACAGAAGTACAGCGCAGATAGATCGGAGAA 778  
Db 1468 GCGCCTTGATCGAGTGCCTGAGGCCGTGAGAAATACAGAGCAGGCTGAGCTCAGAGAG 1527  
QY 779 CAGCCCATACCTGAACCCCTCAGCTGTGTTCAAGCCAGCCAAAGCCATATACAGAATGT 838  
Db 1528 CAGCCCATACCTGAGCTTACAAATTTCTCCACCTGCTAAAGCCATTGACCAAGATTGT 1587  
QY 839 CTCACATTTGTTGGTGGCTGAACCGAGAGAATCTATGCCATGCTGACCCCTACTGTCCC 898  
Db 1588 CTCATACCTACTGTGTGGCTGAGCCGAGCAAGCTCTATGCCATGCTCTCCCTGTATGCC 1647  
QY 899 CGACAGTGAATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGTGT 958  
Db 1648 TGAGGGGACATCAAGGCCCTGACCACTCTCTGTGACCTGGCAGACCGAGAGCTTGTGT 1707  
QY 959 TATCATTTGATGGGCGAAGCATATTCCAGGCTTCTCCAGGCTGTCCCTGGCGACAGAT 1018  
Db 1708 CATCATTTGGCTGGGCCAAGCACAATCCAGGCTTCTCAAGCCTTCTCCCTGGGGACAGAT 1767  
QY 1019 GAGCCTTCTGACAGAGTGTGTTGAATTTTGAATCCTTGTGTGTTATACCGTCTCT 1078  
Db 1768 GAGCCTGCTGACAGAGTGCCTGATGGAATCCTCATCCTGGGATCTGTACCGCTCGCT 1827  
QY 1079 TTCAATTTGAGGATGAACCTGTCTATGACAGACGATTTATATGAAGAGACCACTCCA 1138  
Db 1828 GCCCTACGACGACCAAGCTGTGTATGCTGAGACTACATCATGAGTGAAGACACTCCCG 1887  
QY 1139 ATTAGAGGCCCTTCTGTATCTAAATATGCTATCTCTGACGCTGTAAAGAAATACAAGAG 1198  
Db 1888 CCTGGGGGGCTGCTGAGCTCTACCGGGCCATCTCTGACGTGTTACGAGGTACAAGAA 1947  
QY 1199 CATGAAGCTGAAAAAAGAAATTTGTCAACCCTCAAGCTATAGCTCTTGTCTAATTGAGA 1258  
Db 1948 GCTCAGGTGAGAAAGAGAGATTGTGACGCTCAAGGCCCTGTGCCCAACTCCGA 2007  
QY 1259 CTCATGCAATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAAGATGTCTTACATGAAGC 1318  
Db 2008 TTCCATGTACATCGAGATCTAGAGGTGTCCAGAAGCTGCAGAGACCTGTGCAGAGGC 2067  
QY 1319 GCTCGAGATTATGAAGCTGGCCAGACATGGAAGACCCCTCGTCGAGCTGGCAAGATGCT 1378  
Db 2068 ACTCGAGACTACGAGCTGAGGCCAGGCCCATGAGGAGCCCTGGAGGACGGCAAGTGTCT 2127  
QY 1379 GATGACACTGCCACTCTCTGAGGACAGACTCTACCAAGGCCGTGACGATTTCTACAAAT 1438  
Db 2128 GCTGACACTGCCGCTGCTGCGGACAGCGGCCCAAGGCCGTGACGACTTCTATAGCGT 2187  
QY 1439 CAACTAGAAAGGCAAAAGTCCCAATGACAAACTTTTGGAAATGTTGAGGCCAAGGT 1498  
Db 2188 CAACTGCAAGGCAAAAGTCCCAATGACAAACTTTCTGAGATGCTGAGGCCAAGGC 2247  
QY 1499 CTGACTAAAAGCTCCCT 1515  
Db 2248 CTGGGCGAGGGCTGACT 2264

RESULT 15  
AAA09800  
ID AAA09800 standard; DNA; 2807 BP.  
XX  
AC AAA09800;  
XX  
DT 05-JUL-2000 (first entry)  
XX  
DE Human nuclear receptor nNR1 gene.  
XX  
KW Human nuclear receptor protein-1; nNR1; physiological function; ss;  
XX cell development and differentiation controller; gene expression.  
OS Homo sapiens.  
XX  
PN US6054295-A.

```
XX 25-APR-2000.
PD
XX
XX 26-AUG-1998; 98US-00141000.
PF
XX 27-AUG-1997; 97US-0057090P.
PR 21-OCT-1997; 97US-0062922P.
PR 19-MAR-1998; 98US-0078633P.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Chen F;
XX
XX WPI: 2000-328352/28.
DR P-PSDB; AAY83822.
XX
XX New polynucleotide encoding human nuclear receptor protein-1 is useful
PT for screening compounds that acts as modulators of cell differentiation,
PT cell development and physiological function.
XX
XX Claim 9; Col 57-58; 58pp; English.
PS
XX
CC This sequence represents the DNA encoding human nuclear receptor protein-
CC 1 (nMRI). The protein is a cell development and differentiation
CC controller which is useful for screening compounds that act as modulators
CC of cell differentiation, cell development and physiological functions or
CC for treating and diagnosing disorders associated with altered expression
CC of the DNA
XX
XX Sequence 2807 BP; 701 A; 777 C; 770 G; 559 T; 0 U; 0 Other;
SQ

Query Match 24.0%; Score 715; DB 3; Length 2807;
Best Local Similarity 71.5%; Pred. No. 1.8e-152;
Matches 956; Conservative 0; Mismatches 375; Indels 6; Gaps 1;

QY 179 AGAGCTCTCTGAGATGTCAAAAGATCGACACATTTGAGCTGTTGCTT 238
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 934 AGGGCTGTGAACAGAGATGTCTCGAGACAGGACCTGGGCTCCAGCTGCGCTT 993
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 239 CATCAAGAGGAACCTTCCAGCCAGCCTCCCTGACGACAGCGTCAACCAACAGAGCC 298
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 994 CATCAAGACTGAGCCGTCCAGCCCGTCTCGGGCATAGATGCCCTCAGCCACCAAGCCC 1053
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 299 TGGTGGCTCTTACAGCGCCAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGG 358
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1054 CAGTGGCTGTCCGACGCCAGCGCGGCTTTGGCCTGGCCCTGGGCAACCCAGCCAAACGG 1113
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 359 ACTTGACTGCGCACTCTCTTAACCTTCTGCTCTCTATCCTGGAGGTAAGTGGCCTGTAG 418
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1114 TCTGGACTGCGCAAC-----CATGTTTGACAGGCGCGGGCTGGAGGCCACCCCATGCGG 1167
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 419 GAAACTGTATGATGACTGCTCCAGACCATTTGTTGAAGATCCCAGAACCAAGTGAATA 478
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1168 CAAGAGCTACGAGGACTGTGCCAGGCAATGAGAGACTCGGCCATCAAGTGCGAGTA 1227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 479 CATGCTCACTCGATGAGCCCAAGAGACTGTGTTAGTGTGTGATCATGCTTCTGGGTA 538
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1228 CATGCTCAAGCCCATCCCAAGCGCTGTGCTCTGTGTGCGGGACATTTGCTCTGCTA 1287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 539 CCACTATGGGGTAGCATCATGTGAAGCTGCAAGGCATTTCTTCAAGAGACATTTCAAG 598
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1288 CCACTACGGCGTGGCTCTCTGCGAGGCTTGCAAGGCTTCTTCAAGAGACTATCCAAGG 1347
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 599 CAATATAGATACAGCTGCCCTGCCAGATGAATGTGAATCACAAAGCGACAGCTAA 658
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1348 GAAATTTGATACAGCTGCCCGCGCCCAACCAAGAGATCAACAAAGGAGCGCAA 1407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 659 ATCTGCCAGAGCTTGCCGCTTCATGAAGTGTAAAGTGGGCTAGCTGAAGAAAGGGGT 718
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1408 GTCTGCAAGGCTGCGCTTCATGAATGCTCAAAAGTGGGATGCTGAAGGAAGTGT 1467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 719 GCGTCTTGACAGAGTACGTGAGGTGGCAGAGTACAAAGCGCAGATAGATGCGGAGAA 778
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 1468 GCGCTTGAATCGAGTGCCTGAGAGCCGTCAGAAATACAAAGCGAGCGCTGAGCTCAGAGAG 1527
QY 779 CAGCCCAATACCTGAACCCCTCAGCTGGTTCAAGCCAGCCAAAGGCAATATAAAGATTGT 838
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1528 CAGCCCAATACCTGAGCTTACAAATTTCTCCACCTGCTAAAGGCAATTTGACCAAGATTGT 1587
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 839 CTCACATTTGTTGGTGGCTGACCGGAGAAAGATCTATGCCATGCTGACCCCTACTGTCCC 898
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1588 CTCATACCTACTGTGTGGCTGAGCCGGAACAAGCTCTATGCCATGCTCTCCCTGTATGCC 1647
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 899 CGACAGTACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACGAGAGTTGTGTGT 958
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1648 TGAGGGGACATCAAGGCCCTGACCACTCTCTGTGACCTGGCAGACCGAGAGCTTGTGT 1707
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 959 TATCATTTGATGGGCGGAAGCATATTCAGGCTTCTCCACGCTGTCTCCGACAGAT 1018
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1708 CATCATTTGGCTGGGCCAAGCACATCCAGGCTTCTCAAGCCTCTCCCTGGGGACAGAT 1767
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1019 GAGCCTTCTGACAGTGTCTTGATGGAATAATTTGATCCTTGTGTGTATACCGGTCTCT 1078
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1768 GAGCCTGTGACAGTGTCTTGATGGAATAATCCTCATCTGCGGCACTGTACCGCTGCT 1827
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1079 TTCAATTTGAGATGAACTTGTCTATGACAGCATATATATATGAAGAGACAGTCCAA 1138
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1828 GCCCTACGACGACAAAGCTGTGTAGCTGAGGACTACATCATGATGAAGAGACTCCCG 1887
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1139 ATTAGAGGCTTCTTGTATCTAATAATGCTATCCTGACGCTGTGTAAGAATACAGAG 1198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1888 CCTGCGGGGCTGCTGTGAGCTCTACCGGGCCATCTCTGACGCTGTGATACAGAGTA 1947
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1199 CATGAGCTGCAAAAAGAGATTTGTCAACCTCAAGCTATAGCTTTGCTAATTCAGA 1258
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1948 GCTCAAGGTGAGAAAGAGAGATTGTGACGCTCAAGGCCCTGGCCCTGCCAACTCCGA 2007
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1259 CTCCATGACATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAAGATGTCTTACATGAAGC 1318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2008 TTCCATGTATCATGAGATCTAGAGCTGTCCAGAAAGCTGACGACTGTCAGAGGC 2067
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1319 GCTGAGAGTTATGAAGCTGGCCAGCATGGAAGAACCTCGTGAAGCTGGCAAGATGCT 1378
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2068 ACTGCAAGACTACGAGCTGAGCCAGGCCATGAGAGGCCCTGAGAGACGGCAAGCTGCT 2127
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1379 GATGACACTGCCACTCTGAGGCAAGCTCTACCAAGCCGTGACGATTTCTACACAT 1438
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2128 GCTGACACTGCCGCTGTGCGGCAAGCGCCGCCAAGGCCGTGACACTTCTATAGCGT 2187
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1439 CAAACTAGAAGGCAAGTCCCAATGCACAAATTTTGTGAAATGTTGGAAGCCCAAGT 1498
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2188 CAAACTGCAAGGCAAAAGTGCCCATGACAAACTCTTCTGAGATGCTGGAAGCCCAAGGC 2247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1499 CTGACTAAAAGCTCCCT 1515
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2248 CTGGGCCAGGGCTGACT 2264
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: August 9, 2004, 23:43:00  
Job time : 1138 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2004, 23:08:43 ; Search time 7329 Seconds  
(without alignments)  
12162.452 Million cell updates/sec

Title: US-10-054-841-3  
Perfect score: 2985  
Sequence: 1 gcgggcccgcagtggtgtgtg.....actggcgccgcgcgttgat 2985

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vtl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1913.8	64.1	4835	11	AK052731	AK052731 Mus muscu
2	1375.4	46.1	1377	29	AY412347	AY412347 Homo sapi
3	1349.6	45.2	1377	29	AY412348	AY412348 Pan trogl
4	1229.8	41.2	1377	29	AY412349	AY412349 Mus muscu

5	823.4	27.6	893	9	AL700159	AL700159 DKFZp686J
6	763	25.6	919	13	BX369768	BX369768 BX369768
7	744.6	24.9	1005	13	BX332672	BX332672 BX332672
8	704	23.6	1296	29	AY409734	AY409734 Homo sapi
9	703.2	23.6	799	12	BG434768	BG434768 602507976
10	696.6	22.9	783	14	CA752448	CA752448 UI-M-FOO-
11	684.4	22.3	4005	11	AK032025	AK032025 Mus muscu
12	679.8	22.8	4297	11	AK052256	AK052256 Mus muscu
13	679	22.7	704	28	B60472	B60472 CIT-HSP-200
14	670.4	22.5	1296	29	AY409736	AY409736 Mus muscu
15	667.6	22.4	780	14	CB989379	CB989379 AGENCOURT
16	657.4	22.0	1296	29	AY409735	AY409735 Pan trogl
17	652.2	21.8	738	14	CB520351	CB520351 UI-M-GIO-
18	651.8	21.8	765	14	CF743115	CF743115 UI-M-GV0-
19	646.8	21.7	754	14	CB527738	CB527738 UI-M-FY0-
20	640.4	21.5	3590	11	BC031205	BC031205 Mus muscu
21	632.4	21.2	774	14	CB989610	CB989610 AGENCOURT
22	631.6	21.2	848	14	CF594270	CF594270 AGENCOURT
23	611.6	20.5	784	13	BU289036	BU289036 603608157
24	594.8	19.9	941	13	BX706173	BX706173 BX706173
25	593.2	19.9	662	14	CB724159	CB724159 UI-M-FY0-
26	591	19.8	611	9	AL712433	AL712433 DKFZp686A
27	576.6	19.3	746	13	BU243544	BU243544 603590708
28	571.2	19.1	757	13	BU251126	BU251126 603403749
29	554.2	18.6	691	13	BU285177	BU285177 603868119
30	548.2	18.4	2364	11	AK044339	AK044339 Mus muscu
31	538	18.0	687	10	BG009643	BG009643 MRA-GN019
32	537.4	18.0	563	9	AL700376	AL700376 DKFZp686I
33	532.6	17.8	771	13	BU394797	BU394797 603806351
34	530.2	17.8	647	13	BU359346	BU359346 603477360
35	527.4	17.7	553	13	BX486063	BX486063 DKFZp686A
36	476.6	16.0	814	14	CF745277	CF745277 UI-M-GV0-
37	459.6	15.4	654	9	AA167702	AA167702 zq38908.S
38	452.8	15.2	494	28	AQ221321	AQ221321 HS_2008.A
39	444	14.9	565	13	BU287387	BU287387 604164988
40	443.4	14.9	3360	11	BC044274	BC044274 Xenopus 1
41	433.6	14.5	689	9	AL636193	AL636193 AL636193
42	430.4	14.4	2268	11	BC039774	BC039774 Mus muscu
43	428	14.3	622	14	W26275	W26275 25c6 Human
44	421.6	14.1	673	13	BX706174	BX706174 BX706174
45	421	14.1	730	14	CF182620	CF182620 UI-M-EY0-

ALIGNMENTS

RESULT 1	AK052731	4835 bp	MRNA	linear	HTC 20-SEP-2003
LOCUS	AK052731				
DEFINITION	Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched library, clone:D630039B07 product:estrogen-related receptor gamma, full insert sequence.				
ACCESSION	AK052731				
VERSION	AK052731.1	GI:26342926			
KEYWORDS	HTC, CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				



REFERENCE	
AUTHORS	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 4835)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiroaka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/. Location/Qualifiers  1. 4835  /organism="Mus musculus" /mol_type="mRNA" /strain="C57Bl/6J" /db_xref="FANTOM_DB:D630039B07" /db_xref="MGI:2422648" /db_xref="taxon:10090" /clone="D630039B07" /tissue_type="kidney" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="0 day neonate" 281..1588 /note="unnamed protein product; estrogen-related receptor gamma (MGD MGI:1347056, GB NW_011935, evidence: BLASTN, 99%, match=1444) putative" /codon_start=1 /protein_id="BAC35120.1" /db_xref="GI:26342927"
FEATURES	
source	
CDS	

ORIGIN		/translation="MSNKRHRHIDSSCSFIXTEPSSPASLIDSVNHSPGGSSDASGS YSSITNGHONGIDSPPLYPSPAPISGSEPRKLYDDCSSTVEDPQTKCEYMLNSMPK RLCLVCGDIASGYHYVASCACAEAFKRTIQNIIEYSCEPATNECETKRRRKSQAC RFMKCLKVGMLEKGVRLDRVRGRQYKRRIDAENSPLYNPQLVOPAKPYNKIVSHL LVAEPEKIYAMPDPVTPDSDIKALITLCLDLADRELVLIIWAKHIFGFSTLSDQMS LLOSAMMELLILGVVYRSISFEDELVYADDYIMDEQSKLAGLLDNLNAILQLYKYYK SMKLEKEEFVTLKAILALANSDSMHIEDVEAVQKLQDLVLEHALQDYEAQUMEDPRAG KMLMTLPLLRQSTKAVQHFNYIKLEKVPMHKLFLEMLEAKV"						
Query Match	64.1%;	Score 1913.8;	DB 11; Length 4835;					
Best Local Similarity	83.5%;	Pred. No. 0;						
Matches 2353;	Conservative 0;	Mismatches 412;	Indels 52; Gaps 14;					
QY	163	CCCTGCAC	TACGAGAAAGAGCTTCTCTGCA	GAAATGTCA	AAACAAGATCGACATTTGATT	222		
Db	249	CAC	TGTGCA	TTGGAAAGGCTTCTCTGCA	GAAATGTCA	AAACAAGATCGACATTTGATT	308	
QY	223	CCAGCT	TTCTCTTCATCA	AGACGGAACCTTCCAG	CCCCAGCTTCC	TGACGGACAGCG	282	
Db	309	CCAGCT	TTCTCTTCATCA	AGACGGAACCTTCCAG	CCCCAGCTTCC	TGACGGACAGCG	368	
QY	283	TCAACCA	CCACAGCCCTGTG	GCTCTTCAGACGCCAG	TGGAGCTTCAAGTCA	ACCATGA	342	
Db	369	TCAACCA	CCACAGCCCTGTG	GCTCTTCAGACGCCAG	TGGAGCTTCAAGTCA	ACCATGA	428	
QY	343	ATGGCCAT	CAGAACGGA	CTTGACTCGCACCTCTTAC	CCCTTCTGCTCCTATC	CTGGGAG	402	
Db	429	ATGGCCAT	CAGAACGGA	CTTGACTCGCACCTCTTAC	CCCTTCTGCTCCTATC	CTGGGAG	488	
QY	403	G	TAGTGGCCTGTCTCAG	AAACTGTATGATGACTG	CTCCAGACCATTTG	TGAAGATCCCC	462	
Db	489	GCAGCGG	CGCTGTCCG	AAACTGTATGATGACTG	CTCCAGACCATTTG	TGAAGATCCCC	548	
QY	463	AGACCAAG	TGTAATACATG	CTCACTCGATGCCCA	AGAGACTGTGTTAG	TGTGTGTG	522	
Db	549	AGACCAAG	TGTAATACATG	CTCACTCGATGCCCA	AGAGACTGTGTTAG	TGTGTGTG	608	
QY	523	ACATCGCT	CTGGGTACCA	CTATGGGTA	GCATCATGTGA	AGCCTGCAAGGCATTTCTTCA	582	
Db	609	ACATCGCT	CTGGGTACCA	CTATGGGTA	GCATCATGTGA	AGCCTGCAAGGCATTTCTTCA	668	
QY	583	AGAGGCA	ATTCAAGGCA	ATATAGATA	CAGCTGCCCTGCCA	CGAATGAATGTGAATCA	642	
Db	669	AGAGGCA	ATTCAAGGCA	ATATAGATA	CAGCTGCCCTGCCA	CGAATGAATGTGAATCA	728	
QY	643	CAAA	CGCAGACG	TAAATCTG	CCAGGCTTGCCG	CTTCAATGAAGTGT	TTAAAGTGGCA	702
Db	729	CAAA	CGCAGACG	TAAATCTG	CCAGGCTTGCCG	CTTCAATGAAGTGT	TTAAAGTGGCA	788
QY	703	TGCTGA	AGAAAGGGGTG	CGTCTTGA	CAGAGTACGTG	AGCTCGGCAGAA	AGTACAAGCCCA	762
Db	789	TGCTGA	AGAAAGGGGTG	CGTCTTGA	CAGAGTACGTG	AGCTCGGCAGAA	AGTACAAGCCCA	848
QY	763	GGATAG	ATGCGGAGAACAG	CCCATACCTGA	ACCCTCAGCTG	GTTCAGCCAGCCAA	AAAAAGC	822
Db	849	GAATAG	ATGCTGA	AACAGCCCATAC	CTGAACCCCTCAG	CTGTGTGCAAGCCAA	AAAAAGC	908
QY	823	CATATA	CAAGATTGTCT	CACATTTGTTG	TGGTGAACCGG	AGAGATCTATG	CCATGC	882
Db	909	CATATA	CAAGATTGTCT	CACATTTGTTG	TGGTGAACCGG	AGAGATCTATG	CCATGC	968
QY	883	CTGACCT	TA	CTGCCCCGACAG	TGACATCA	AAAGCCCTCACTA	CACTGTGTGACTTGGCCG	942
Db	969	CTGACCT	TA	CTGCCCCGACAG	TGACATCA	AAAGCCCTCACTA	CACTGTGTGACTTGGCCG	1028
QY	943	ACCGAG	ATTGGTGTATCA	TTGGATGGGCGA	AGCATATTCCAG	GCTTCTCCAGCCTGT	1002	
Db	1029	ACCGAG	ATTGGTGTATCA	TTGGATGGGCGA	AGCATATTCCAG	GCTTCTCCAGCCTGT	1088	
QY	1003	CCCTG	CGGACCA	GATGAGCCTTCTG	CAGAGTGC	TTGGAATTTTGA	TCTTGGTG	1062
Db	1089	CCCTG	CGGACCA	GATGAGCCTTCTG	CAGAGTGC	TTGGAATTTTGA	TCTTGGTG	1148

QY	1063	TCGTATACCGGTCTCTTTTCATTTTGAGATGAACCTGTCTATGCAACGATTATATATG	1122
Db	1149	TTGTGTACCAATCGCTTTGCGTTTGAGATGAACCTGTCTATGCAACGATTATATATG	1208
QY	1123	ACGAAGACCAGTCCAAATTTAGCAGGCTCTTGATCTAAATATATGCTATCCGACGTGG	1182
Db	1209	ATGAAGACCAGTCTAAATTTAGCAGGCTCTTGACCTAAATATATGCTATCCGACGTGG	1268
QY	1183	TAAAGAAATACAAGACATGAAGCTGAAAAAGAAATTTGTCAACCTCCAAAGCTATAG	1242
Db	1269	TGAAGAAGTACAAGACATGAAGCTGAGAAGAAAGAAATTCGTCAACCTCAAGCAATAG	1328
QY	1243	CTCTTGCTAATTCAGACTCCATGCGACATAGAAGATGTTGAAGCCGTTCAAGACTTCAGG	1302
Db	1329	CTCTTGCTAATTCAGATTCATGCAATATGAAAGATGTGAAGCTGTGCAGAAACTTCAGG	1388
QY	1303	ATGTCTTACATGAAGCGCTGCAGGATTAAGAAGCTGGCCAGCACATGGAAGACCCCTGTC	1362
Db	1389	ATGTGTACATGAAGCCCTGCAGATTAAGAAGCTGGCCAGCACATGGAAGACCCCTGCC	1448
QY	1363	GAGCTGGCAAGATGCTGATGACACTGCCACTCCTGAGGCAGACCTTAACCAAGCCGTGC	1422
Db	1449	GTGCAGGCACAGATGCTGATGACGCTGCCGCTGCTGAGGCAGACCTCAACCAAGCAGTGC	1508
QY	1423	AGCAATTTCTACACATCAAACTAGAAGGCAAAGTCCCAATGCACAACTTTTGGAAA	1482
Db	1509	AGCACTTCTACACATCAAACTGGAAGGCAAAGTCCCATGCACAACTTTTGGAAA	1568
QY	1483	TGTTGGAGGCCAAGGTCTGACTAAAAG-CTCCCTGGGCCCTTCCCATCTTCATGTTGA	1540
Db	1569	TGCTGAGGCCAAGGTCTGACTAAAAGCCCCCTGGGCCCTCCCATCTGCAAGTTG-A	1627
QY	1541	AAAGGAAATATAAACCAAGAAGTGTGGAAGAACTTAGAGTTTAGTTAACACATCA	1600
Db	1628	AAAGGAAGATATAAACCAAGAATGATGTGAAGATCTTAGAGTTTAGTTAACACATTA	1687
QY	1601	AAAAATCAACAGACTGCACTGATAATTTAGCAGCAAGACTATGAAGCAGCTTTCAGATTCC	1660
Db	1688	AAAAATCAACAGACTGCACTGATATTTTAGCAGCCACAGTACGATGCAGCCGCGGATTCC	1747
QY	1661	TCCATAGGTTCCCTGATGAGTTCTTTCTACTTCT-CCATCATCTTCTTCTCTTCTTC	1719
Db	1748	GCTACATCTTCTGATAGGTTCTCTCTACTTTATCCACGATCTCTGGCCACATCCCTG	1807
QY	1720	CCACATTTCTCTTCTCTTTAATTTTCTCCCTTTCTCTTTCACCTCCCTTAATTTCTT	1779
Db	1808	CATTCCTCCACTCTTCTTGTCTATATTATATGTTGGCTTCTTTCACTAATAGTTTCAAT-	1866
QY	1780	GCTTCTTCATTCCTAGTTCCCAATCTCCTTATTTCTTCCCGTGCCTGCTCTT	1839
Db	1867	-----TTTCCCTCTCCCGTCTCTCTCTCTCCCFCCCTCTCTGTCTCCCCCTTCTT	1921
QY	1840	CTTTTCTTACCTACTCTTCATTTCTCTCTTTTCTTCATCTTCCCTTTTTCATAATTG	1899
Db	1922	CTTTTCTCTCTTCTTCCACAACTTCTCTCTCTTGCTTGCTCTCACCTCTCTTGCGCTTC	1981
QY	1900	AAATAGCTTTAGTTTAAAAAAAATCTCTCCCTTCCCTTCTTCTT--CCCTTCTT	1956
Db	1982	TCACATCTCTCCCACTCTGCGTACATAGTCAATACCTTGATTTGATGAACATTTCTT	2041
QY	1957	TCCTTTTCCTTCTCTCTTTCCTT--TTCCTTTCCTTCTCTTGAACCTTCTTCCATC	2013
Db	2042	TTACCTCTTGCACTCTCTCTCGCTCTCTTCTTCCCACTTTTGTGTTGTTGTTG	2101
QY	2014	TTTTCTTTTCTCTCTCTGCTGTAACTTTAAAAAGAGTCTTAACG-----	2063
Db	2102	TTTTCTTTCTCTCTCTCTGCTGTAACCTTAATAGCAGTCTTAACGAGAGAGAAA	2161
QY	2064	--AAGAGAGATGGAAGCCAGCCCTGCCAAAGAGATGAAGATCATAATATGATGCCAGTG	2121
Db	2162	GAGAGAGAGATGGAAGCCAGCCCTGCCAAAGAGACAGAGATCCATATATGATGCCAGTG	2221

QY	2122	AACTTATGTGAACCATACCGTCCCCAATGACTAAGGAATCAAGAAGAACCAACGCT	2181
Dp	2222	AACTTGTCAATGAACCATGACATCCCAAGTAGTAAGGAATCAAGAGAG-----AACCG	2275
QY	2182	TCCTAAAAAGTACAGTGCACAATATACAATTTGACTGAGTGCAGTATTGATTTCAATGGGA	2241
Dp	2276	TACCTAAGATACATTGCAACGCAACGAGATCAACTTAGTGCAGTATTGATTCTACCGG-	2334
QY	2242	GCAGCCTCTAATTAGACAACCTTAAGCAACGTGCATCGGCTGCTTCTTAATCATTTGCTTTT	2301
Dp	2335	GCAGCCTTCGATCAGACAAACCTAA-----GTGGCGGCAATGGCTGCTCTCTGCTTTC	2389
QY	2302	CCATCTAGATCAGTTTACAGCCATTGATTCCTTAATGTTTTCAGTCTTCCAGGTAT	2361
Dp	2390	TCATCTAGATCAGTTTACAGCCATTGATTCCTTAATGTTTTCAGTCTTCCAGGTAT	2449
QY	2362	TTGTTAGTTTACGTAATAATGTAATTTTTCAGGGAATAGTTTAAAGCTTTATTCATTATG	2421
Dp	2450	TGGTTAGTTTACGTAATAATGTAATTTTTCAGGGAATCCTTTAAGCTTTATTCATTATG	2509
QY	2422	CAATTAATAAGAGAAATGAATTAAGTACTGCAATTTTGCTGCTGCTTTGAACAATTAACA	2481
Dp	2510	CAATTAATAAGAGAGAGGATGAAGATACCGCAACCTGCTGCTTGAACAATTT--GAACA	2567
QY	2482	ATAATGAAGACAAATGAATCCTGGAAGGAAGATTTTAAAAATGTTTGTCTTCTTAC	2541
Dp	2568	CTAATGAAGACAAATGAACCTGGAAGGAAGATTTTAAAAATGTTTGTCTTCTTAC	2627
QY	2542	AAATGAGATTTTGTGTAACAGCTTTTACCACATTTTTCAGCCATTATTAATTAAGGAATT	2601
Dp	2628	AAATGAGATTTTGTGTAACAGCTTTTACCACATTTTTCAGCCATTATTAATTAAGGAATT	2687
QY	2602	TAACTTACTCAAGCAATAGTTGAAGGGAAGGTGCATATTATCAAGGATGCAATTTATGTT	2661
Dp	2688	TAACTTACTCAAGCAATAGTTGAAGGGAAGGTGCATATTATCAAGGATGCAATTTATGTT	2747
QY	2662	GTTGTCCAGTCTGTTCCCAACAATCAATTTCTTAAATGATGAGCTCCAGTTTACCTAAATGT	2721
Dp	2748	GTTGTCCAGTCTGTTCCCAACAATCAATTTCTTAAATGATGAGCTCCAGTTTACCTAAATGT	2806
QY	2722	TCACTGACACAAAGGATGAGATTACACCTACAGTACTTGAGTAGTCAATATATAAGC	2781
Dp	2807	TCACTGACACAAAGGATGAGATTACACCTGCGGAGACCGAGTGTCCCATCCAGAGC	2866
QY	2782	ACTGCAATGAGATATAGATCCGTAAGATTGTACAGAGTGCACTCTCTACTTTGGGAGGT	2841
Dp	2867	ACTGCAATGAGATCCCTATCTGTAGAATTAGCACCACTACACCTCCCTGC-CGGGAGGG	2925
QY	2842	ACAAATTGCCATATGATTTCTAGCTGCCATGTTGTTTGAAGATGTGATACTGCTGTTGC	2901
Dp	2926	ACAGTGCACATACGTTTCTAGCTGCCCTCGTGTATGGAACAAAGATGCTGCTGTATAC	2985
QY	2902	AAAGTACAGACCTTGCCTCAGAGAAGCTGTGAGCCAGTATTCATTTAAGAGATT	2958
Dp	2986	AAACTC-----TGTCTCAGAGAAGAGCTGTGAGCCATATCAATTTCAGAGGCAAT	3034

RESULT 2	
AY412347	
LOCUS	
DEFINITION	AY412347      1377 bp    DNA       linear   GSS 16-DEC-2003 Homo sapiens ESRRG gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION	AY412347
VERSION	AY412347.1     GI:39768312
KEYWORDS	GSS.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1377) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.U., Sninsky,J.U.,
AUTHORS	



TITLE Adams,M.D. and Cargill,M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 1377)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Fierliera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..1377  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>1377  
/gene="ESRRG"  
/locus\_tag="HCM4527"  
ORIGIN

Query Match 46.1%; Score 1375.4; DB 29; Length 1377;  
Best Local Similarity 99.9%; Pred. No. 9.3e-275;  
Matches 1376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 126 ATGATTCGGTAGAACTTTGCTTCCGATCTTTTCCCTGCAGTACGAGAGAAGCTT 185  
Db 1 ATGATTCGGTAGAACTTTGCTTCCGATCTTTTCCCTGCAGTACGAGAGAAGCTT 60  
QY 186 CTCTGCAAGATGTCAAAACAAGATGCACACATTTGTCAGCTGTTCTCTTCAACAAG 245  
Db 61 CTCTGCAAGATGTCAAAACAAGATGCACACATTTGTCAGCTGTTCTCTTCAACAAG 120  
QY 246 ACGGAACCTTCCAGCCCCAGCCCTCCCTGACGGAAGCGTCAACCAACCAAGCCCTGTGCG 305  
Db 121 ACGGAACCTTCCAGCCCCAGCCCTCCCTGACGGAAGCGTCAACCAACCAAGCCCTGTGCG 180  
QY 306 TCTTCAGACGCCAGTGGAGCTACAGTTCAACCATGAATGGCCATCAAGACGACTTGAC 365  
Db 181 TCTTCAGACGCCAGTGGAGCTACAGTTCAACCATGAATGGCCATCAAGACGACTTGAC 240  
QY 366 TCGCCACCTCTTACCCTTCTGCTCCATCTCTGGAGGTAGTGGGCTGTCAAGAACTG 425  
Db 241 TCGCCACCTCTTACCCTTCTGCTCCATCTCTGGAGGTAGTGGGCTGTCAAGAACTG 300  
QY 426 TATGATGACTGTCCAGCACCATTGTGTAAGATCCCCAGACCAAGTGTGAATACATGCTC 485  
Db 301 TATGATGACTGTCCAGCACCATTGTGTAAGATCCCCAGACCAAGTGTGAATACATGCTC 360  
QY 486 AACTCGATGCCCAAGAGACTGTGTTAGTGTGTGACATCGCTTCTGGGTACCACTAT 545  
Db 361 AACTCGATGCCCAAGAGACTGTGTTAGTGTGTGACATCGCTTCTGGGTACCACTAT 420  
QY 546 GGGGTAGATCATGTGAAGCCTGCAAGGCATTCTTCAAGAGACAATTCAAGGCAATATA 605  
Db 421 GGGGTAGATCATGTGAAGCCTGCAAGGCATTCTTCAAGAGACAATTCAAGGCAATATA 480  
QY 606 GAATACAGCTGCCCTGCCACGAATGATGTGAATCAAAAAGCGCAGACGTAATCCTGC 665  
Db 481 GAATACAGCTGCCCTGCCACGAATGATGTGAATCAAAAAGCGCAGACGTAATCCTGC 540  
QY 666 CAGGCTGCCGCTTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGGGTGCGCTTT 725  
Db 541 CAGGCTGCCGCTTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGGGTGCGCTTT 600  
QY 726 GACAGAGTACGTGGAGTCCGCAGAAATACAAAGCGCAGATAGATCGGAGAACAGCCCA 785  
Db 601 GACAGAGTACGTGGAGTCCGCAGAAATACAAAGCGCAGATAGATCGGAGAACAGCCCA 660

QY 786 TACCTGAACCCCTCAGCTGGTTCAGCCAGCCAAAAGCCATATAACAAGATTGTTCAAT 845  
Db 661 TACCTGAACCCCTCAGCTGGTTCAGCCAGCCAAAAGCCATATAACAAGATTGTTCAAT 720  
QY 846 TTGTTGGTGGCTGAACCCGAGAAAGATCTATGCCATGCTGACCTACTGTCCCGCAGT 905  
Db 721 TTGTTGGTGGCTGAACCCGAGAAAGATCTATGCCATGCTGACCTACTGTCCCGCAGT 780  
QY 906 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACGAGAGTTGTTGTTACATT 965  
Db 781 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACGAGAGTTGTTGTTACATT 840  
QY 966 GGATGGCGAAGCATATTCCAGGCTTCTCCACGCTGTCCCTGGCGGACAGATGAGCCTT 1025  
Db 841 GGATGGCGAAGCATATTCCAGGCTTCTCCACGCTGTCCCTGGCGGACAGATGAGCCTT 900  
QY 1026 CTGCAGAGTCTTGGATGGAATTTTGATCCTTGGTGTGCTATACCCGTCTCTTCAATT 1085  
Db 901 CTGCAGAGTCTTGGATGGAATTTTGATCCTTGGTGTGCTATACCCGTCTCTTCAATT 960  
QY 1086 GAGGATGAACCTGTCTATGACAGACGATTATATATGGAAGCAAGACCAAGTCCAAATTAGCA 1145  
Db 961 GAGGATGAACCTGTCTATGACAGACGATTATATATGGAAGCAAGACCAAGTCCAAATTAGCA 1020  
QY 1146 GGCCCTCTTGATCTAAATTAATGCTATCTCTGACGCTGTTAAAGAAATACAGAGCATGAAG 1205  
Db 1021 GGCCCTCTTGATCTAAATTAATGCTATCTCTGACGCTGTTAAAGAAATACAGAGCATGAAG 1080  
QY 1206 CTGGAAGAAAGAAATTTGTCAACCCCTCAAGCTATAGCTCTTGTAAATTCAGACTCCATG 1265  
Db 1081 CTGGAAGAAAGAAATTTGTCAACCCCTCAAGCTATAGCTCTTGTAAATTCAGACTCCATG 1140  
QY 1266 CACATAGAAGATGTGAAGCCGTTCAAGAGCTTCAAGATGCTTACATGAAGCGCTGACG 1325  
Db 1141 CACATAGAAGATGTGAAGCCGTTCAAGAGCTTCAAGATGCTTACATGAAGCGCTGACG 1200  
QY 1326 GATTATGAAGCTGGCCAAGCATGGAAGCCCTGTCAGAGCTGGCAAGTCTGATGACA 1385  
Db 1201 GATTATGAAGCTGGCCAAGCATGGAAGCCCTGTCAGAGCTGGCAAGTCTGATGACA 1260  
QY 1386 CTGCCACTCTGAGGCGACAGCCTTACCAAGGCCGTGCAGCATTTCTACACATCAAACTA 1445  
Db 1261 CTGCCACTCTGAGGCGACAGCCTTACCAAGGCCGTGCAGCATTTCTACACATCAAACTA 1320  
QY 1446 GAAGGCAAGTCCCAATGCACAACTTTTTTGGAAATGTTGGAGGCCAAGGTCTGA 1502  
Db 1321 GAAGGCAAGTCCCAATGCACAACTTTTTTGGAAATGTTGGAGGCCAAGGTCTGA 1377

RESULT 3  
AY412348 1377 bp DNA linear GSS 16-DEC-2003  
LOCUS  
DEFINITION Pan troglodytes ESRRG gene, VIRUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY412348  
VERSION AY412348.1 GI:39768313  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
REFERENCE 1 (bases 1 to 1377)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Fierliera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 1377)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,



Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.

COMMENT

FEATURES  
source  
1.1377  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>1377  
/gene="ESRRG"  
/locus\_tag="HCM4527"

ORIGIN

Query Match 45.2%; Score 1349.6; DB 29; Length 1377;  
Best Local Similarity 98.4%; Pred. No. 2.1e-269;  
Matches 1355; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 126 ATGATTCGGTAGAAGCTTTCCTCTCTGAATCTTTTCCCTGCACTACGAGGAAGAGCTT 185  
Db 1 ATGATTCGGTAGAAGCTTTCCTCTCTGAATCTTTTCCCTGCACTACGAGGAAGAGCTT 60  
QY 186 CTCGACAGATGTCAACAAGAATCGACACATTTGATTCAGCTGTTCTCTCATCAAG 245  
Db 61 CTCGACAGAAATGTCAACAAGAATCGACACATTTGATTCAGCTGTTCTCTCATCAAG 120  
QY 246 ACGGAACCTTCCAGCCCGAGCCCTCCCTGACCGGACAGCGTCAACGACAGCCCTGCTGCG 305  
Db 121 ACGGAACCTTCCAGCCCGAGCCCTCCCTGACCGGACAGCGTCAACGACAGCCCTGCTGCG 180  
QY 306 TCTTCAGACGCCAGTGGAGCTTACAGTTCAACCATGAATGGCCATCAGAACGGAAGCTTAC 365  
Db 181 TCGTCAGACGCCAGTGGAGCTTACAGTTCAACCATGAATGGCCATCAGAACGGAAGCTTAC 240  
QY 366 TCGCCACTCTCTACCCCTTCTCTCTATCTCTGGAGGTAGTGGCCCTGTCAAGAACTG 425  
Db 241 TCGNNNNNTCTTACCCCTTCTCTCTATCTCTGGAGGTAGTGGCCCTGTCAAGAACTG 300  
QY 426 TATGATGACTGTCTCCAGACCAATTTGTTGAAGATCCCCAGACCAAGTGTGAATACATGCTC 485  
Db 301 TATGATGACTGTCTCCAGACCAATTTGTTGAAGATCCCCAGACCAAGTGTGAATACATGCTC 360  
QY 486 AACTCGATGCCCAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGTACCACATAT 545  
Db 361 AACTCGATGCCCAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGTACCACATAT 420  
QY 546 GGGGTAGCATCATGTGAAGCCTGCAAGGATTTCTCAAGAGGACAAATTCGAAGGCAATATA 605  
Db 421 GGGGTAGCGTCATGTGAAGCCTGCAAGGCAATTTCTCAAGAGGACAAATTCGAAGGCAATATA 480  
QY 606 GAATACAGCTGCCCCGTCACGAATGAATGAATCACAAGGCGCAGACGTAAATCCTGCG 665  
Db 481 GAATACAGCTGCCCCGTCACGAATGAATGAATCACAAGGCGCAGACGTAAATCCTGCG 540  
QY 666 CAGGCTTGCCTTCATGAAGTGTAAAAAGTGGCATGCTGAAGAAGGGGTGCGTCTT 725  
Db 541 CAGGCTTGCCTTCATGAAGTGTAAAAAGTGGCATGCTGAAGAAGGGGTGCGTCTT 600  
QY 726 GACAGAGTACGTGAGGTGCGCAGAGTCAAGCGCAGATAGATGCGGAGAAACAGCCCA 785  
Db 601 GACAGAGTACGTGAGGTGCGCAGAGTCAAGCGCAGATAGATGCGGAGAAACAGCCCA 660  
QY 786 TACCTGAACCTCAGCTGTTTACAGCCAGCCAAAAGCCATATATACAGATTTGTTACAT 845  
Db 661 TACCTGAACCTCAGCTGTTTACAGCCAGCCAAAAGCCATATATACAGATTTGTTACAT 720  
QY 846 TTGTTGGTGGTGAACCGGAGAGATCTATGCCATGCTGAGCCCTACTGTCCCGACAGT 905  
Db 721 TTGTTGGTGGTGAACCGGAGAGATCTATGCCATGCTGAGCCCTACTGTCCCGACAGT 780

QY 906 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGGTGTATCATTT 965  
Db 781 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGGTGTATCATTT 840  
QY 966 GGATGGCGGAAGCATATTCAGGCTTCTCCACGCTGTCCCTGGCGGACCATGAGCCCTT 1025  
Db 841 GGATGGCGGAAGCATATTCAGGCTTCTCCACGCTGTCCCTGGCGGACCATGAGCCCTT 900  
QY 1026 CTGCAAGAGTGTGGATGGAATTTTGATCTCTTGGTGTGCTATACCGGTCTCTTTCATTT 1085  
Db 901 CTGCAAGAGTGTGGATGGAATTTTGATCTCTTGGTGTGCTATACCGGTCTCTTTCGTTT 960  
QY 1086 GAGGATGAACCTGTCTATGCAAGCATTTATATATATGACGAGAACCCAGTCCAAATTAGCA 1145  
Db 961 GAGGATGAACCTGTCTATGCAAGCATTTATATATATGACGAGAACCCAGTCCAAATTAGCA 1020  
QY 1146 GGCCTTCTGATCTAAATATGCTATCTCTGACGCTGGTAAAGAAATACAAAGAGCATGAAG 1205  
Db 1021 GGCCTTCTGATCTAAATATGCTATCTCTGACGCTGGTAAAGAAATACAAAGAGCATGAAG 1080  
QY 1206 CTGGA AAAAGAGAAATTTGTCAACCTCAAGCTATAGCTCTGCTAATTTCAGACTCCATG 1265  
Db 1081 CTGGA AAAAGAGAAATTTGTCAACCTCAAGCTATAGCTCTGCTAATTTCAGACTCCATG 1140  
QY 1266 CACATAGAAGATGTTGAAGCCGTTCAGAAGCTTCAGAGATGTCTTACATGAAGCGCTGCA 1325  
Db 1141 CACATAGAAGATGTTGAAGCTGTTCAGAAGCTTCAGAGATGTCTTACATGAAGCGCTGCA 1200  
QY 1326 GATTATGAAGCTGGCCAGCACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGATGACA 1385  
Db 1201 GATTATGAAGCTGGCCAGCACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGATGACA 1260  
QY 1386 CTGCCACTCCTGAGGAGAGACCTCTACCAAGGCGGTGAGCATTTCTACAAATCAAACTA 1445  
Db 1261 CTGCCACTCCTGAGGAGAGACCTCTACCAAGGCGGTGAGCATTTCTACAAATCAAACTA 1320  
QY 1446 GAAAGCAAGTCCCAATGACAAACTTTTTTGGAAATGTTGAGGCCCAAGGTCTGA 1502  
Db 1321 GAAAGCAAGTCCCAATGACAAACTTTTTTGGAAATGTTGAGGCCCAAGGTCTGA 1377

RESULT 4  
AY412349 1377 bp DNA linear GSS 16-DEC-2003  
LOCUS  
DEFINITION Mus musculus ESRRG gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY412349  
VERSION AY412349.1 GI:39768314  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
Location/Qualifiers

```

source
1. .1377
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1. .>1377
/gene="ESRRG"
/locus_tag="HCM4527"

ORIGIN

Query Match      41.2%;   Score 1229.8;   DB 29;   Length 1377;
Best Local Similarity 93.3%;   Pred. No. 1.5e-24;
Matches 1285;   Conservative 0;   Mismatches 92;   Indels 0;   Gaps 0;

QY      126  ATGATTCGGTAGAACTTTGCGCTTCTGAAATCTTTTCCCTGCACCTAGAGGAGAGCTT 185
Db      1  ATGATTCGGTAGAACTTTGCGCTTCTGAAATCTTTTCCCTGCACCTAGAGGAGAGCTT 60

QY      186  CTCTGCAGAAATGTCAAAACAAGATCGACATGATTCCAGCTGTTGCTTCTCATCAAG 245
Db      61  CTCTGCAGAAATGTCAAAACAAGATCGACATGATTCCAGCTGTTGCTTCTCATCAAG 120

QY      246  ACGGAACCTTCCAGCCCGACCTCCCTGACGGAACAGCGTCAACCAACAGCCCTGTGGC 305
Db      121  ACGGAACCTTCCAGCCCGACGCTCCCTGACGGAACAGCGTCAACCAACAGCCCTGTGGG 180

QY      306  TCTTCAGAGCCAGTGGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGACTTGAC 365
Db      181  TCTTCCGACGCCAGTGGGAGTTACAGTTCAACCATGAATGGCCATCAGAACGGACTGGAC 240

QY      366  TCGCCACCTTCTTAACCTTCTGCTCTTATCTCTGGGAGGTAGTGGGCTGTCAAGAACTG 425
Db      241  TCGCCACCTTCTTAACCTTCTGCTCTTATCTCTGGGAGGTAGTGGGCTGTCCGAAACTG 300

QY      426  TATGATGACTGCTCCAGCACCATTTGTTGAAGATCCCCAGAACCAAGTGTGAATACATGCTC 485
Db      301  TATGATGACTGCTCCAGCACCATCTGTAAGAGATCCCGACCAAGTGTGAATATATGCTC 360

QY      486  AACTCGATGCCCAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACCACTAT 545
Db      361  AACTCGATGCCCAAGAGACTGTGTTAGTGTGTGTGCGACATCGCTTCTGGGTACCACTAT 420

QY      546  GGGGTAGCATCATGTGAAGCCTGCAAGCATTTCTCAAGAGACAATTCAAGGCATATA 605
Db      421  GGGGTGATCATGTGAAGCCTGCAAGCATTTCTCAAGAGACAATTCAAGGTACATA 480

QY      606  GAATACAGCTGCCCTGCCACGAATGAATGTGAATCACAAAGCCGACAGCTAAATCTCTGC 665
Db      481  GAGTACAGCTGCCCGACGACGAATGAATGTGAGATCACAAAGCCGACAGCAAAATCTCTGC 540

QY      666  CAGGCTTGGCGCTTTCATGAAGTGTAAAAAGTGGGCATGCTGAAGAAGGGGTGCGTCTT 725
Db      541  CAGGCTTGGCGCTTTCATGAAGTGTTCAAAGTGGGCATGTGAAGAAGGGGTCCGTCCTT 600

QY      726  GACAGAGTACGTGAGGTGGGACAGAGTACAAGCGCAGAGATAGTGGGAGAACAGCCCA 785
Db      601  GACAGAGTGCCTGAGGTGGGACAGAGTACAAGCGCAGAGATAGTCTGAGAACAGCCCA 660

QY      786  TACCTGAACCCCTCAGCTGTTCAAGCCAGCCAAAAAGCCCATATAAAGATTGTCTACAT 845
Db      661  TACCTGAACCCCTCAGCTGTTCAAGCCAGCCAAAAAGCCCATATAAAGATTGTCTCGCAT 720

QY      846  TTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGACCTACTGTCCCGCAGAGT 905
Db      721  TTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGACCTACTGTCCCGCAGAGT 780

QY      906  GACATCAAGCCCTCAGTACATGTGTGACTTTGGCCGACCGAGAGTTGGTGTATCATTT 965
Db      781  GACATCAAGCCCTCAGCACAACACTGTGTGACTTTGGCTGACCGAGAGTTGGTGTATCATTT 840

QY      966  GGAATGGGCGAAGCATATTTCCAGGCTTCTCCAGCGCTGTCCCTGGCGGAGACAGATGAGCCTT 1025
Db      841  GGAATGGGCGAAGCATATTTCCAGGCTTCTCCAGCACTGTCCCTGGCGAGACAGATGAGCCTTC 900
```

```

QY      1026  CTGCAGAGTCTTGATGAGAAATTTGATCCTTGGTGTGCTATACCGGCTCTTTCATTT 1085
Db      901  CTCCAGAGTGCATGATGAGATTTGATCCTCGGCGTGTGTACCGATCGCTTTCGTTT 960

QY      1086  GAGATGAAGTGTCTATGACAGATTAATATATGAGACGAAGACCAAGTCCAATTAGCA 1145
Db      961  GAGATGAAGTGTCTATGACAGATTAATATATATGATGAGACGACGCTAAATTAGCA 1020

QY      1146  GGCTTCTTGAATCTAATATATGCTATCTCGACAGCTGGTTAAGAAATACAGACATGAAG 1205
Db      1021  GGCTTCTTGAACCTTAATATATGCTATCTCGACAGCTGGTTAAGAAATACAGACATGAAG 1080

QY      1206  CTGGAAGAAAGAGAAATTTGTCAACCCCTCAAGCTATAGCTCTTGTCTAAATTCAGACTCCATG 1265
Db      1081  CTAGAGAGAGAGAAATTCGTCAACCCCTCAAGCAATAGCTCTTGTCTAAATTCAGATTCATG 1140

QY      1266  CACATAGAAAGATGTTGAAGCCGTTCAAGAGCTTCAGAGATGTCTTAATGAAGCGCTGCAG 1325
Db      1141  CATATAGAAAGATGTTGAAGCTGTGCAAGAACTTCAGAGATGTCTTAATGAAGCGCTGCAG 1200

QY      1326  GATTATGAAGCTGGCCAGCACATGGAAGACCTCGTCAGCTGGCAAGATGCTGATGACA 1385
Db      1201  GATTACGAGGCTGGCCAGCACATGGAAGACCTCGCGCTGCAAGCAAGATGCTGATGACG 1260

QY      1386  CTGCCACTCCTGAGGCGACAGCCTTCAACAGCCCGTGCAGCATTTCTACACATCAACTA 1445
Db      1261  CTGCCGCTGCTGAGGCGACAGCCTCCACCAAGCAGTCCAGCACTTCTACACATCAACTC 1320

QY      1446  GAAGGCAAGTCCCAATGCACAACTTTTTTTGGAATGTGTGAGGCCCAAGCTCTGA 1502
Db      1321  GAAGGCAAGTCCCATGCACAACTTTTTTTGGAATGTGTGAGGCCCAAGGTCTGA 1377
```

```

RESULT 5
AL700159
LOCUS
DEFINITION
DKFZp686J14117 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION
AL700159
VERSION
AL700159.1 GI:19620692
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 893)
Ottenwaelder, B., Obermaier, B., Mewes, W., Mewes, H.W., Weil, B. and
Wieman, S.
EST (Ottenwaelder, B., Obermaier, B., Mewes, H.W., Weil, B. and
Wieman, S.)
Unpublished (2001)
COMMENT
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wieman, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wieman@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFZp686J14117) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
```

FEATURES

```

source
1. .893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J14117"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1tb="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
```

ORIGIN		cdna-collection"
Query Match	27.6%;	Score 823.4; DB 9; Length 893;
Best Local Similarity	98.7%;	Pred. No. 3e-160;
Matches	883;	Conservative 0; Mismatches 6; Indels 6; Gaps 5;
QY	1447	AAGCAAGTCCCAATGCACAACCTTTTGTGAAATGTTGGAGGCCAAGTCTGACTAA 1506
Db	1	AAGCAAAGTCCCAATGCACAACCTTTTGTGAAATGTTGGAGGCCAAGTCTGACTAA 60
QY	1507	AAGTCCCCTGCGCCTTCCCATTCCCTTCATGTTGAATAAAGGAAAATAAACCAAGTGTAT 1566
Db	61	AAGTCCCCTGCGCCTTCCCATTCCCTTCATGTTGAATAAAGGAAAATAAACCAAGTGTAT 120
QY	1567	GTCGAAGAAGCTTAGAGTTTAGTTTAACAACATCAAAAAATCAACAGACTGCATGATAATT 1626
Db	121	GTCGAAGAAGCTTAGAGTTTAGTTTAACAACATCAAAAAATCAACAGACTGCATGATAATT 180
QY	1627	TAGCAGCAAGACTATGAAGCAGCTTTCAGATTCCCTCCATAGGTTCCGTATGAG-TTCTTT 1685
Db	181	TAGCAGCAAGACTATGAAGCAGCTTTCAGATTCCCTCCATAGGTTCCGTATGAGTTCTTT 240
QY	1686	CTACTTCTCGCATCATCTTCTTTCCCTCTTTCTTCCCACATTTCTCTTCTTATTTTTT 1745
Db	241	CTACTTCTCGCATCATCTTCTTTCCCTCTTTCTTCCCACATTTCTCTTCTTATTTTTT 300
QY	1746	TCTCCTTTTCTTCTTTCACCTCCCTTATTTCTTGTCTTCTTTCATTCCTAGTCCCATTC 1805
Db	301	TCTCCTTTTCTTCTTTCACCTCCCTTATTTCTTGTCTTCTTTCATTCCTAGTCCCATTC 360
QY	1806	TCCCTTATTTTCTTCCGCTCTGCTGACCCTTCTTCTTCTTCTTACCTACTCATTCCTC 1865
Db	361	TCCCTTATTTTCTTCCGCTCTGCTGACCCTTCTTCTTCTTCTTACCTACTCATTCCTC 420
QY	1866	TCTTTTCTCATCCTTCCCTTTTCTTCTAAATTGGAATAGCTTTAGTTTAAAAAATAAAA 1925
Db	421	TCTTTTCTCATCCTTCCCTTTTCTTCTAAATTGGAATAGCTTTAGTTT-AAAAAAATA 478
QY	1926	TCCTCCCTTCCCCCTTTCCTTTCCCTTTCTTTCCTTTTCCCTTTTCCCTTTCCCT 1985
Db	479	TCCTCCCTTCCCCCTTTCCTTTTCCCTTTCTTTCCTTTTCCCTTTTCCCTTTCCCT 538
QY	1986	TTCTTTTCTCTTGACCTTTTCCATCTTTCTTCTTCTTCTTCTTCTGCTGTAACCTTT 2045
Db	539	TTCTTTTCTCTTGACCTTTTCCATCTTTCTTCTTCTTCTTCTTCTGCTGTAACCTTT 598
QY	2046	AAAAAGGCTCTTAACCTGAAGAGATGGAAGCCAGCCCTGCCAAGGAT-GGAGATCCA 2104
Db	599	AAAAAGGCTCTTAACCTGAAGAGATGGAAGCCAGCCCTGCCAAGGATGGAGATCCA 658
QY	2105	TAAATATGATGCCAGTGAACCTTATTTGTGAACCATACCGTCCCAATGACTAAGGAATCAA 2164
Db	659	TAAATATGATGCCAGTGAACCTTATTTGTGAACCATACCGTCCCAATGACTAAGGAATCAA 718
QY	2165	AGAGAGAGAACCACGTTCCCTAATAAAGTACAGTGCAACATATACAATTGACTGAGT-GCA 2223
Db	719	AGAGAGAGAACCACGTTCCCTAATAAAGTACAGTGCAACATATACAATTGACTGAGTGGCA 778
QY	2224	GTAATTAGATTTCAT-GGGAGCAGCCTCTAATTAGACAACCTTAAAGCAACGTTGCATCGGCT 2282
Db	779	GTAATTAGATTTCATGGGAGCAGCCTCTAATTAGACAACCTTAAAGCAACGTTGCATCGGCT 838
QY	2283	GCTTCTTATCATTTGCTTTTCCATCTAGATCAGTTACAGCCATTGATTCTTAAT 2337
Db	839	GCTTCTTATCATTTGCTTTTCCATCTAATTCAGTTACAGCCATTGATTCTTAAT 893
RESULT 6		
LOCUS	BX369768	919 bp mRNA linear EST 08-MAY-2003
DEFINITION	BX369768 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens CDNA clone CS0DC014YD14 5-PRIME, mRNA sequence.	

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
EX369768	1	GI:30441500	EST.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished (2001)	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 904.r For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAG007ZG02_CS00591_1&cluster=904.r. Contact : Feng liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0BAG007ZG02_CS00591_1.
FEATURES	source	1. .919	/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"	/clone="CS0DC014YD14"	/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"	/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"	/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	Query Match	25.6%;	Score 763;	DB 13;	Length 919;				
	Best Local Similarity	98.6%;	Pred. No. 1e-147;						
	Matches 799;	Conservative 0;	Mismatches 8;	Indels 3;	Gaps 3;				
QY	2145	CCCAATGACTAAGAAATCAAGAAGAGAGAACCAAGCTTCTAAAGTACAGTGCACATA							2204
DB	36	CCCCAAGACTAAGAAATCAAGNAGAAGAACCAACGTTCT-AAAGTACAGTGCACATA							94
QY	2205	TACAAATTGACTGAGTGCAGTATTAGATTTCATGGGAGCAGCCCTTAATTAGACAATTAA							2264
DB	95	TACAAATTGACTGAGTGCAGTATTAGATTTCATGGGAGCAGCCCTTAATTAGACAATTAA							154
QY	2265	AGCAACGTTGCATCGGCTGCTTCTATCATTCATTGCTTTCCATCTAGATCAGTTACAGCCAT							2324
DB	155	AGCAACGTTGCATCGGCTGCTTCTATCATTCATTGCTTTCCATCTAGATCAGTTACAGCCAT							214
QY	2325	TTGATTCCTTAATTGTTTTTTTCAAGCTCTCCAGGTATTTGTTAGTTAGTACTATGTAA							2384
DB	215	TTGATTCCTTAATTGTTTTTTTCAAGCTCTCCAGGTATTTGTTAGTTAGTACTATGTAA							274
QY	2385	CTTTTTCAGGGAATAGTTTAAGCTTTATTCATTCATGCATACTAAAGAGAAATTAAGAT							2444
DB	275	CTTTTTCAGGGAATAGTTTAAGCTTTATTCATTCATGCATACTAAAGAGAAATTAAGAT							334
QY	2445	ACTGCAATTTTGTGCTGCTTGAACAATTAAGCAACAATAAAGAGCAAAATGAATCCT							2504
DB	335	ACTGCAATTTTGTGCTGCTTGAACAATTAAGCAACAATAAAGAGCAAAATGAATCCT							394
QY	2505	GAAGGAAGATTTTAAAAATGTTTGTCTTCTTACAAATGGAGATTTTGTGACAG							2564
DB	395	GAAGGAAGATTTTAAAAATGTTTGTCTTCTTACAAATGGAGATTTTGTGACAG							454
QY	2565	CTTTACCACTTTTCAGCCATTTAATAATGGGAATTAACCTTCAAGCAATAGTTGA							2624
DB	455	CTTTACCACTTTTCAGCCATTTAATAATGGGAATTAACCTTCAAGCAATAGTTGA							514



QY	2625	AGGGAAGGTGCATATTTATCAGGATGCAATTATGTTGTGCCAGTCTGTGTCCTCCAAACA	2684
Db	515	AGGGAAGGTGCATATTTATCAGGATGCAATTATGTTGTGTGCCAGTCTGTGTCCTCCAAACA	574
QY	2685	TCAATTTCTTAACATGAGCTCCAGTTTACCCTAAATGTTCACTGCACAAAGATGAGATT	2744
Db	575	TCAATTTCTTAACATGAGCTCCAGTTTACCCTAAATGTTCACTGCACANAGATGAGATT	634
QY	2745	ACACCTACAGTGAAGCTGTGAGTAGTCACATATATATAAGCACTGCACATGAGATATAGATCCG	2804
Db	635	ACACCTACAGTGAAGCTGTGAGTAGTCACATATATATAAGCACTGCACATGAGATATAGATCCG	694
QY	2805	TAGAATTGTCAAGAGTGCACCTCTCTACTTGGGAGGTACAAATTGCCATATGATTTTCTAGC	2864
Db	695	TAGAATTGTCAAGAGTGCACCTCTCTACTTGGGAGGTACAAATTGCCATATGATTTTCTAGC	754
QY	2865	TGCCATGTGTTAGGAATGTGATTAAGTGCCTGTGTTGGCAAAGTCAACAGACCTTGCTTCAGA	2924
Db	755	TGCCATGTGTTAGGAATGTGATTAAGTGCCTGTGTTGGCAAAGTCAACAGACCTTGCTTCAGA	812
QY	2925	AGGAGCTGTGAGCCAGTATTTCAATTTAAGAG	2954
Db	813	AGGAGCTGTGAGCCAGTATTTCAATTTTAAAG	842

RESULT 7	
BX332672	
LOCUS	
DEFINITION	
BX332672	1005 bp
BX332672 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens	mRNA linear EST 01-MAY-2003
cDNA clone CS0DC014YD14 5-PRIME, mRNA sequence.	

ACCESSION	BX332672	Accession number, unique sequence.
VERSION	BX332672.1	GI:30312142
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 1005)	
TITLE	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.	
JOURNAL	Full-length cDNA libraries and normalization	
COMMENT	Unpublished (2001)	
	Contact: Genoscope	

Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: [segregf@genoscope.cns.fr](mailto:segregf@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 904.r For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DC014DB07QPl&cluster=904.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DC014DB07QPl&cluster=904.r). Contact :  
Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
[http://fulllength.invitrogen.com/Invitrogen Corporation 1600](http://fulllength.invitrogen.com/InvitrogenCorporation1600)  
Faraday Avenue Genoscope sequence ID : CS0DC014DB07QPl.

```

FEATURES
source
location/Qualifiers
1..1005
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC014YD14"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

```
ORIGIN
      600 bp contig 8 vector library was normalized.
Query Match          24.9%; Score 744.6; DB 13; Length 1005;
Best Local Similarity 88.6%; Pred. No. 6.6e-14;
Matches 820; Conservative 43; Mismatches 53; Indels 9; Gaps 6;
QY    1640 ATGAAGCAGCTTCAGATTCCCTCCATAGTTCCTGATGAG-TTCTTTACTTTCCCAT 1698
|||||
```

Db	64	ATGAAGCAGCCTTTCAGATTCCTCCATAGAGTTCCTGATGAGTTTCTTCTACTTCTCCAT	123
QY	1699	CATCTTCTTTCCCTCTTTCTTTCCACACATTTCTCTTTCTTTATTTTCTCTCCCTTTCTTC	1758
Db	124	CATCTTCTTTCTCTTTCTTTCCACACATTTCTCTTTCTTTATTTTCTCTCTCTTTCTTC	183
QY	1759	TTTCACTTCCCTTATTTCTTTGCTTCTTTTCATTCGAGTTCCCATTTCTCTTATTTTCT	1818
Db	184	TTTCACTTCCCTTATTTCTTTGCTTCTTTTCATTTCCGATTTCCCATTTCTCTTATTTTCT	243
QY	1819	TCCCGTCTGCCTGACCTTCTTTCTTTTCTTTTACCTACCTCATTTCCCTCTCTTTCTCATCC	1878
Db	244	TCCCGTCTGCCTGACCTTCTTTCTTTTCTTTTACCTACCTCATTTCCCTCTCTTTCTCATCC	303
QY	1879	TTCCCTCTTTTCTTAAATTGAAATAGCTTTAGTTTAAAAAAAATCCCTCCCTTCCCTC	1938
Db	304	TTCCCTCTTTTCTTAAATTGAAAMARCTTTAGTTT - AAAAAAATCCCTCCCTTCCCTC	361
QY	1939	CTTTCCTTTCCCTTTCTTTCTCTTTTCTTCCCTTTCCCTTTTCCCTTTCCCTTTCTCTT	1998
Db	362	CTTTCCTTTCCCTTTCTTTCTCTTTTCTTCCCTTTCCCTTTTCCCTTTCCCTTTCTCTT	421
QY	1999	GACCTTCTTTCCATCTTTCTTTTCTTCTCTCTCTGCTGTAACCTTTTAAAGAGTCTCT	2058
Db	422	GACCTTCTTTCCATCTTTCTTTTCTTCTCTCTCTGCTGTAACCTTTTAAARAGTCTCT	481
QY	2059	AACTGAAGAGATGGAAGCCAGCCCTGCCAAAGATGAGATCCATAATATGATGCCA	2118
Db	482	AACTGAAGAGATGGAAGCCAGCCCTGCCAAAGATGAGATCCATAATATGATGCCA	541
QY	2119	GTGAAC-TTATTTGTGAACCATACCGTCCCAATGACTAAGGAATCAAGAGAGAGACCA	2177
Db	542	GTGAACATTATTTGTGAACCAWACCGTCCCAATGACTAAGGAATCAAGAGAGAGACCA	601
QY	2178	ACGTTCCTAAAAAGTACAGTGCACACATATATACAAATTTGACTGAGTGCAGTATTGATTTCA	2237
Db	602	ACGTTCCTAAAAAGTACAGTGCACACATATATACAAATTTGACTGAGTGCAGTATTGATTTCA	661
QY	2238	GGGAGCAGCCTCTAATTAGACAACCTTAAGCAACGTTCGCTGCTCTTATCATTC	2297
Db	662	GGGAGCAGCCTCTAATTARAMAAMTTAAGCAAGTTGCATYGSCTGCTCTTATCATTC	721
QY	2298	TTTTCCATCTAGATCAGTTACAGCCATTTGATTCCTTAATTTGTTTTTCAAGTCTTCCAG	2357
Db	722	TTTTCCATCTAGATCAGTTACAGCCATTTGATTCCTTAATTTGTTTTTCAAGTCTTCCAG	781
QY	2358	GTAATTTGTTAGTTAGCTACTATGTAATTTTTCAGGGAATAGTTTAAAGCTTATTCATTT	2417
Db	782	GTAATTTGTTAGCTACTATGTAATTTTTCAGGGAATAGTTTAAAGCTTATTCATTT	841
QY	2418	CATGCAATCTTAAAGAGAAATAGAATACCTGCAATTTTGTGCTGCTTTGAACAAATTACG	2477
Db	842	TAKCSAA-AATTAAGMRMAAAAAAATAAAGCAA-WTTCGCTGCTTTTTCAAAAATTAAG	899
QY	2478	AACAATATGAAGCACAATGAATCCTGGAAGAGAGATTTTAAAAATGTTTGTCTTC	2537
Db	900	CAAAATATATGARGCAAAAWGATTTWKACGSCABATTTT - AAAAKYTTTXXTTTTTT	956
QY	2538	TTACAAATGAGATTTTTTTTGTACC 2562	
Db	957	TTAAAAAKGCGMWTTTTTTXXTTAAAC 981	

RESULT	8
AY409734	
LOCUS	
DEFINITION	AY409734 1296 bp DNA linear GSS 16-DEC-2003 Homo sapiens ESRB gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION	AY409734
VERSION	AY409734.1 GI:39765702
KEYWORDS	GSS.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

[illegible]

Db	535	CGTGAAGGCCGCTCAGAAATACAAGCGACGGGCTGGACTCAGAGAGCAGCCCATACCTGAGC	594
QY	795	CCTCAGCTGEGTTACGCCAGCCAAAAAGCCATATACACAGATTGTTCACTTTGTTGTTG	854
Db	595	TTACAAATTTCCTCCACCTGTCTAAAAAGCCATTGACCAAGATTGTTCTATCCTACTGTTG	654
QY	855	GCTGAACCGGAGAAAGATCTATGCCATGCGCTGACCCCTACTGTGCCGACGTACATCAAA	914
Db	655	GCTGAGCCCGACAAAGCTCTATGCCATGCGCTCCCTGTGATGCTGAGGGGGAACATCAAG	714
QY	915	GCCCTCACTACACTGTGTGACTTTGGCCGACGAGAGTTGTTGTTATCATTTGATGGCG	974
Db	715	GCCCTGACCACTCTCTGTGACCTGGCAGACCGAGAGCTTGTGTCATCATTTGGCTGGCC	774
QY	975	AAGCATATTCAGAGCTTCTCCACGCTGTCCCTGGCCGACCCAGATGAGCCTTCTGACAGT	1034
Db	775	AAGCACATCCCGAGCTTCTCAAGCCTCTCCCTGGGGGACCCAGATGAGCCTGTGACAGT	834
QY	1035	GCTTGATGGAATTTTGATTCCTTGGTGTGCTATACCCGCTCTCTTTCATTTGAGATGAA	1094
Db	835	GCCTGATGGAATTCCTCATCTCTGGGCATCGTGTATCCGCTCGCTGCCCTACGACGACAAG	894
QY	1095	CTTGTCATATGACAGCATTTATATATATGACGAGAACCAAGTCCAAATTGACGAGCTTCTT	1154
Db	895	CTGTTGATCGCTGAGGACTACATCATGATGAGAGGACACTCCCGCCTCGCGGGGCTGCTG	954
QY	1155	GATCTAATAATATGCTATCTCTGACAGCTGGTAAAGAAATACAAGAGCATGAAGCTGGAAAA	1214
Db	955	GAGCTCTACCGGGCCATCTCTGACGTGGTAGCGGATACAAGAAAGCTCAAGGTGAGAGAG	1014
QY	1215	GAAGAAATTGTACCCCTCAAAAGCTATAGCTCTTGCTAATTCAGACTCCATGCACATAGAA	1274
Db	1015	GAGGAGTTTGTGACGCTCAAGGCCCTGGCCCTCGCCCACTCCGATTCCATGTACATCGAG	1074
QY	1275	GATGTTGAAGCCGTTCAAGAGCTTCAGGATGTCTTACATGAAGCGCTGCAGAGATTATGAA	1334
Db	1075	GATCTAGAGGCTGTCCAGAAGCTGCAGGACCTGTCTGCACGAGGCACTGCAGGACTACGAG	1134
QY	1335	GCTGGCCAGCACATGGAAGACCCCTGTCGAGCTGGCAAGATGCTGATGACACTGCCACTC	1394
Db	1135	CTGAGCCAGCGCCCATGAGAGAGCCCTGGAGGACGGGCAAGCTGCTGTGACACTGCCGCTG	1194
QY	1395	CTGAGGCAAGACCTCTACCAAGGCCGTGCAGCATTTCTACAACATCAAACTAGAGGCANA	1454
Db	1195	CTGCGGCAAGACGGCCGCCCAAGGCCGTGCAGGACCTTCTATAGCGTCAAACTGCAGGGCAA	1254
QY	1455	GTCCCAATGCACAAACTTTTTTTGGAAATGTTGAGGCCAAG	1496
Db	1255	GTGCCATGACAAACTTCTTCTGTGAGATGCTGAGGGCCAAAG	1296

RESULT 9  
BG434768  
LOCUS  
DEFINITION  
799 bp mRNA linear EST 14-MAR-2001  
602507976F1 NIH\_MGC\_79 Homo sapiens cDNA clone IMAGE:460494 5',  
mRNA sequence.  
BG434768  
ACCESSION  
BG434768.1 GI:13341274  
VERSION  
EST.  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 799)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: CLONTECH Laboratories, Inc.  
cDNA library Preparation: CLONTECH Laboratories, Inc.  
CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
COMMENT  
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1CM1341 row: h column: 19  
High quality sequence stop: 627.  
Location/Qualifiers

FEATURES  
source

1. .799  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4604994"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH MGC 79"  
/note="Organ: placenta; Vector: pDNR-LIB (Clontech); site\_1: Sfil (ggcgccgcggcc); site\_2: Sfil (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATCTAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 23.6%; Score 703.2; DB 12; Length 799;  
Best Local Similarity 96.0%; Pred. No. 2.6e-135;  
Matches 753; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

QY	1996	CTTGACCTCTTCTCCATCTTCTTTCTCTCTCTCTGCTGTAACCTTTAAAGAGGTC	2055
Db	1	CTTGACCTCTTCTCCATCTTCTTTCTCTCTCTCTGCTGTAACCTTTAAAGAGGTC	60
QY	2056	TCTAACTGAAGAGAGATGGAAGCCAGCCCTGCCAAAGAGATGAGATCCATAATATGATG	2115
Db	61	TCTAACTGAAGAGAGATGGAAGCCAGCCCTGCCAAAGAGATGAGATCCATAATATGATG	120
QY	2116	CCAGTGAATTAATGTGGAACCATACCGTCCCAATGACTAAGGAATCAAAGAGAGAGAAC	2175
Db	121	CCAGTGAATTAATGTGGAACCATACCGTCCCAATGACTAAGGAATCAAAGAGAGAGAAC	180
QY	2176	CAACGTTCTTAAAGTACAGTGCACATATACAAATTGACTGAGTGCACTATTAGATTTC	2235
Db	181	CAACGTTCTTAAAGTACAGTGCACATATACAAATTGACTGAGTGCACTATTAGATTTC	240
QY	2236	ATGGAGCAGCCTCTAATTAGACAACCTTAAAGCAACGTTGCATCGGCTTCTTATCAT	2295
Db	241	ATGGAGCAGCCTCTAATTAGACAACCTTAAAGCAACGTTGCATCGGCTTCTTATCAT	300
QY	2296	GCTTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTAATGTTTTCACAGTCTCC	2355
Db	301	GCTTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTAATGTTTTCACAGTCTCC	360
QY	2356	AGGTATTGTAGTTAGCTACATGTAACCTTTTCAGGGAATAGTTTAAGCTTATTC	2415
Db	361	AGGTATTGTAGTTAGCTACATGTAACCTTTTCAGGGAATAGTTTAAGCTTATTC	420
QY	2416	TTCAATGCAATACTAAGAGAAATAGAATACTGCAATTTTGCTGCTTGAACAATTA	2475
Db	421	TTCAATGCAATACTAAGAGAAATAGAATACTGCAATTTTGCTGCTTGAACAATTA	480
QY	2476	CGAACATAATGAAGGACAAATGAATCTGAAGGAAGATTTTAAAAATGTTTGTCT	2535
Db	481	CGAACATAATGAAGGACAAATGAATCTGAAGGAAGATTTTAAAAATGTTTGTCT	540
QY	2536	TCTTACAAATGGAGATTTTGTGACAGCTTTACCACTTTTCAGCCATTATTAATAG	2595
Db	541	TCTTACAAATGGAGATTTTGTGACAGCTTTACCACTTTTCAGCCATTATTAATAG	600
QY	2596	GGAATTTAATCTCAAGCAATAGTTGAAGGAAGGTGCATATATCAGG-GATGCAT	2654
Db	601	GGAATTTAATCTCAAGCAATAGTTGAAGGAAGGTGCATATATCAGGCTTGCAT	660

QY	2655	TTATGTTGTGCGCAGTCTGTGTCGCCAAACATCAATTTCTTACATGAGCTCCAGTTACC	2714
Db	661	TTATGTTGTGCGCAGTCTGTGTCGCCAAACATCAATTTCTTACATGAGCTCCAG-TTACC	719
QY	2715	TAAATGTTACTGACACAAAGGATGAGATTACACCTACAGTGACTGTGAGTGCACATA	2774
Db	720	TAAATGTTACTGAC-CCAAAGGTGCGATTACACCTACAGGAGTGTGAGTGCATTT	778
QY	2775	TATA	2778
Db	779	TTTA	782

RESULT 10  
CA752448  
LOCUS  
DEFINITION  
UI-M-FOO-cdo-f-13-0-UI.r1 NIH\_BMAP\_F00 Mus musculus cDNA clone  
IMAGE:6831086 5', mRNA sequence.  
CA752448  
CA752448.1 GI:25583085  
EST.  
Mus musculus (house mouse)  
Mus musculus  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 783)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.  
Location/Qualifiers

FEATURES  
source

1. .783  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6831086"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 12.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH BMAP\_F00"  
/note="Organ: Brain; Vector: pYX-Asc; Site\_1: EcoR I; Site\_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAGAGAGCC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 23.3%; Score 696.6; DB 14; Length 783;  
Best Local Similarity 93.1%; Pred. No. 6.1e-134;  
Matches 729; Conservative 0; Mismatches 54; Indels 0; Gaps 0;



QY	358	GACTTGACTCGCCACCTCTCTACCCCTTCTGCTCCTATCCTGGAGGTAGTGGCCTGTCA	417
Db	1	GACTGGACTCGCCACCTCTCTA	60
QY	418	GGAACCTGTATGATGACTGCTCCAGCACCATTTGTAAGATCCCAGACCAAGTGAAT	477
Db	61	GGAACCTGTATGATGACTGCTCCAGCACCATCGTAGAGATCCCAGACCAAGTGAAT	120
QY	478	ACATGCTCACTCGATGCCCAAGAGACTGTGTTAGTGTGTGGTGACATCGCTTCTGGGT	537
Db	121	ATATGCTCACTCCATGCCCAAGAGACTGTGCTTAGTGTGTGGCGACATCGCCTCTGGGT	180
QY	538	ACCACCTATGGGGTAGCATCATGTGAAGCCTGCAAGGCATTTCTCAAGAGACATTCAG	597
Db	181	ACCACCTATGGGGTAGCATCATGTGAAGCCTGCAAGGCATTTCTCAAGAGACATTCAG	240
QY	598	GCAATATAGATACAGCTGCCCTGCCACGATGAATGAATCACAAGCGCAGACGTA	657
Db	241	GTAACATAGAGTACAGCTGCCACGACGAATGAATGATGATCACAAGCGCAGACGCA	300
QY	658	AATCCTGCCAGGCTTGCCGCTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGCG	717
Db	301	AATCCTGCCAGGCTTGCCGCTCATGAAGTGTCAAAAGTGGGCATGCTGAAGAAGGCG	360
QY	718	TGCGTCTTGACAGAGTAGTGGAGTCCGCAGAGTACAGCGCAGATAGATGCGGAGA	777
Db	361	TCCGTCTTGACAGAGTAGTGGAGTCCGCAGAGTACAGCGCAGATAGATGCTGAGA	420
QY	778	ACAGCCCATACCTGAACCCCTGAGTGTTCAGCCAGCCAAAAGCCATATACAAGATTG	837
Db	421	ACAGCCCATACCTGAACCCCTGAGTGTGACAGCCAAAAGCCATATACAAGATTG	480
QY	838	TCTCACAATTTGTTGGTGGCTGAACCGGAGAAGATCTATGCCATGCTGACCTACTGTCC	897
Db	481	TCTCACAATTTGTTGGTGGCTGAACCGGAGAAGATCTATGCCATGCTGACCTACTGTCC	540
QY	898	CCGACAGTGAACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGATTGGTG	957
Db	541	CCGACAGTGAACATCAAAAGCCCTCACTACACTGTGTGACTTGGCTGACCGAGATTGGTG	600
QY	958	TTATCATTTGGATGGCGGAAGCATATTCAGGCTTCTCCAGCTGTCCCTGGCGAGCCAGA	1017
Db	601	TTATCATTTGGATGGCGGAAGCATATTCAGGCTTCTCCAGCTGTCCCTGGCGAGCCAGA	660
QY	1018	TGAGCCTTCTGCAGAGTGTGATGGAATTTTGATCCTTGCTGTATATACCGGTCTC	1077
Db	661	TGAGCCTTCTGCAGAGTGTGATGGAATTTTGATCCTTGCTGTATATACCGGTCTC	720
QY	1078	TTTCATTTGAGGATGAACCTGTCTATGACAGACGATTATATTAATGACGAAGACCATCCCA	1137
Db	721	TTTCATTTGAGGATGAACCTGTCTATGACAGACGATTATATTAATGATGAAGACCATCTA	780
QY	1138	AAT 1140	
Db	781	ATT 783	

RESULT 11  
AK032025  
LOCUS  
DEFINITION  
AK032025 4005 bp mRNA linear HTC 18-SEP-2003  
Mus musculus adult male medulla oblongata cDNA, RIKEN full-length  
enriched library, clone:6330549E01 product:estrogen related  
receptor, beta, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
AK032025  
AK032025.1 GI:26327830  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Carninci, P. and Hayashizaki, Y.

TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253 10349636 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374 11042159 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 120530913 11076861 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 4005) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:htp://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:htp://genome.gsc.riken.go.jp/ URL:htp://fantom.gsc.riken.go.jp/. location/Qualifiers 1. .4005 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:6330549E01" /db_xref="MGI:2395414" /db_xref="taxon:10090"
---------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

```

/clone="6330549E01"
/sex="male"
/tissue_type="medulla oblongata"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
143. 144

```

```
/note="unnamed protein product; estrogen related receptor,
beta (MGI:MGI:1346832, GB|NM_011934, evidence: BLASTN,
99%, match=1315)
putative"
```

```

/codon_start=1
/protein_id="BAC27656..1"
/db_xref="GI:26327831"
/translation="MSSEDRLHSSCGSFIKTEPSSPSGIDALSHHSPPGSGSDASGGG
FGIALSTHANGLDSPPMFAGAGGPNPCRSKYEDCTGIMEDSAIKCEYMLNAIPKRL
CLVCGDITASGYHYGVASCEACAKFFKRTIQGNIEYNCDATNECEITKRRRSCQACRF
MKCLKYGMLKEGVRLDRVRGGRQKYYKRLDSENSPYNLPISPPAKKPLTKIVSNLLG
VEQDKIAMPNPNDIPEGDIKALTLCIADRLGLVFLINAKHIIGFPSLTLDQMSLL
QSAMWEILLIGIVRSLPYDDKLAAYADVIMDEEHSRLVLDLYRAILOVRRYKKL
KVEKEEMILKALALANDSDSMTIENLEAVOKLQDLLHVALQDYELSQRHEEPRAAGKL
LTLTFLRLQTAQAVQHFYSVYKQGVPMHKLFLIEMLEAKV"
3984..3989
polyA signal

```

```
polyA_site 4005 /note="putative"
/note="putative"
```

Query Match	22.9%	Score 684.4;	DB 11;	Length 4005;
Best Local Similarity	70.0%	Pred. No. 1.8e-131;		
Matches 938; Conservative	0;	Mismatches 396;	Indels 6;	Gaps 1;

Dy      163 CCCTGCACTACGAAGAGAAGCTTCTCTGCAGAATGTCAAACAAGATCGAACACTTTGATT    222  
         |||||  
Db      111 CCTTGCTACCACAACCAAGCTGCTGAACCGAATGTCTTCGGAAAGACAGGCCACTTGGCT    170

2y 223 CCAGCTGTTTCGTCCCTTATCAAGACGGAACTTCCAGCCTCCCTGCACCGACAGCG 282  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 171 CTAGTTCGGCTCTCTTATCAAGACGGAGCCATCCAGCCCCCTCGGGCATTTGATGCC 230

Dy 283 TCAACCAACCAGCCCTGTGGCTTTCAGACGCCAGTGGGAAGCTACAGTTCAACCATGA 342

Dd 231 TCAGCACCACAGCCCCAGCGGCTGTCGAGACGCCAAGTGTCCTTGGCAATTGCCTGA 290

OY		343 ATGGCCATAGAACGGACTTGACTGCCACCTCTCAACCCTTGTGTCCTATCCTGGAG	402
D6		291 GCACCCACGCACAAGGTTCTGAACTGCCCGCTATTGTTCCAGGTGC-----GGGGTTGG	344

403 GTAGTGGCCCTGACGAACTGATGATGACTGCTCCAGCACCATTGTTGAAGATCCC 462  
 345 GAGGCAACCCGTGCCCAAGAGCTACGAGGACTGTACTAGTGTATCATGAGGACTCCG 404

Db 405 CCATCAATGCGAGTACATGCTTAACGCCATCCCAAGCGCTGTGCCTCGTGTGCGGG 464

465 ACATTGCTCTGGCTACCACTACGAGTGGCTCTCTGCGAGGCTTGCAAGGCTTCTTCA 524

Db 525 AGAGAACCAATTCAGGCAACATCGAGTACAACCTGCCCGCCACCAATGAATGTGAGATCA 584

Db 585 CCAACGAGCGCAAGTCTGTCAAGCCTTCGGATTATGAATGCCCTCAAAGTGGGA 644

645 TGCTAAGGAAGTGTGCGCCTTGACCGAGTTGAGGAGGCCCGCAGAACTCAACCGAC 704

[illegible]

Db 705 GGCTGATTCGAGACAGCCCCCTACCTGAACCTGCCGATTCCCCACCTGCTAAAAAGC 764

```

QY      823 CATATAACAGATGTCTCAATTGTTGGTGCCTGAACCGGAGAAGTCTATGCCATGC   882
          ||| |||| | | | | | | | | | | | | | | | | | | | | | |
Db      765 CATTGACTAAGATCGTCTGCAATCTACTAGGGGTGAGCAGACAAGCTGTATGCTATGC   824

```

Dy 883 CTGACCCCTACTGTCCCCCGAGTGCATCAAAGGCCCTCACTACACTGTGTGACTTGCCCG 942

Dd 825 CTCCCAACGATATCCCCCGAGGAGATATCAAGGCCCTGACCACCTCTCTGTGAATTGGCAG 884

943 ACCGAGATGGTGGTATCATGGATGGCGGAAGCATATCCAGGCTTCCACGCTGT 100  
 944  
 Db 885 ATCGGGGCGTTGTCTTCATCACTGGGCCAAGCACATCCGAGCTTCCCCAGTCTGA 944

Db 945 CACTTGGGACCCAGATGAGCCTGTGCAGAGTGCTGATGGACATTCTCATCTTGGCA 100

Db 1005 TCGTACCGCTCGCTCCATACGATGACAAAGCTGGCATAACGCCGAGACTATATCATGG 106

Db 1065 ATGAGGACACTCTGCGCTGTAGGGCTGCTGGACCTTACCGAGCCATCCTGCAGCTGG 1122

Db 1125 TGGCAGGTACAGAACTCAAGGTAGAGAGGAGAGTTTATGATCCTCAAGGCCCTGG 118

Db 1185 CCCTCGCCAACTCAGATTGATGTACATTGAGAACCTGGAGGGGTGCAGAAGCTCCAGG 124

Db 1245 ACCGTGTCACGAGGCGCTGCAGGACTATGAGCTGAGTCAAGGCCACGAGGAGCCCGGA 130

Dd  
1305 GGGCCGCGCAAGCTGTGCTGTACGCCTGCCCTTGCTGAAGCACAAGCCGCCAAGAAGCCGTGC

Dv  
1423 AGCATTTTCTACAAATTCATAATTAGGAAAAAGTGGGAAAATCCCCAACACTTTTTTTTTTCCGAAA

Db 1365 AACACTTACAGTGTGAACTGCAGGGCAAGTGCCCATGCACAAACTCTTCCGTGAGA 1442

Db 1425 TGCTGAGGCCAAGGTGTGA 1444

AK052256	LOCUS	DEFINITION
AK052256	4297 bp	mRNA
Mus musculus	13 days embryo heart	linear
		HTC 20-SEP-2003
		cdna, RIKEN full-length enriched

Accession	Full insert sequence.
AK052256	
AK052256.1	GI:26342471

SOURCE ORGANISM	
Mus musculus (house mouse)	
Mus musculus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
VOLUME  
NUMBER  
PAGES  
DATE

1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
with 5' RACE  
Nucleic Acids Res. 1995  
23: 1088-1092

REFERENCE	2	Cavanagh D	Chibata Y	Hasegawa Y	Saitoh Y	Shibata Y
AUTHORS						
99279253						
MEDLINE						
10349636						
PUBMED						

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes  
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)  
PUBMED 20499374  
REFERENCE 11042159  
AUTHORS 3  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
TITLE RIKEN integrated sequence analysis (RISA) system-384-format  
JOURNAL sequencing pipeline with 384 multicapillary sequencer  
MEDLINE Genome Res. 10 (11), 1757-1771 (2000)  
PUBMED 20530913  
REFERENCE 11076861  
AUTHORS 4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE 11076861  
AUTHORS 5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
TITLE Analysis of the mouse transcriptome based on functional annotation  
JOURNAL of 60,770 full-length cDNAs  
MEDLINE Nature 420, 563-573 (2002)  
PUBMED 12042159  
REFERENCE 11076861  
AUTHORS 6  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirozane,T.,  
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Takeeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
COMMENT cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/.  
FEATURES  
source location/Qualifiers  
1. 4297  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:D330014B12"  
/db\_xref="MGI:2421188"  
/db\_xref="taxon:10090"  
/clone="D330014B12"  
/tissue\_type="heart"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="13 days embryo"  
438. 1739  
/note="unnamed protein product; estrogen related receptor,  
beta (MGD|MGI:1346832, GB|NM\_011934, evidence: BLASTN,

99%, match=1315)  
putative"  
/codon\_start=1  
/protein\_id="BAC34898.1"  
/db\_xref="GI:26342472"  
/translation="MSSEDRHLGSSCGSFIKTEPSSPSGIDALSHSPSGSSDASGG  
EGIALSTHANGIDSPMFAGAGLGGNPGCRKSYEDCTSGIMEDSAIKERYMLNAPKRL  
CLVCGDIASGYHGVASCEACKAFKRTIOGNIENPCATNECEITKRRKSCQACRF  
MKCLKVGMLKEGVRLDRVRGRQKRYKRLDSNSPYLNLPISPPAKPLTKIVSNLIG  
VEQDKLYAMPENDIPEGDIKALTLLCELDRELVFLINWAKHIGFPPSLDQMSLL  
QSAWMEILIGIVYRSLPYDDKLAYAEIYIMDEHSRLVGLDLXRAIIQLVRYKKL  
KVEKEEFMILKALALANSDSMTIENLEAVQKLQDLHEALQDYELSGRHEPRRAKGL  
LTLPLRLQTAKAQVHFYSVKLQKVPMEKLFLEMLEAKV"  
polyA\_signal 4280. 4285  
/note="putative"  
polyA\_site 4297  
/note="putative"  
ORIGIN  
Query Match 22.8%; Score 679.8; DB 11; Length 4297;  
Best Local Similarity 70.2%; Pred. No. 1.6e-130;  
Matches 928; Conservative 0; Mismatches 387; Indels 6; Gaps 1;  
QY 182 GCTTCTCTGCAGAAATGTCAAACAAAGATGCACACATTTGATTCAGCTGTTCTGCTTCAT 241  
Db 425 GCTGCTGAACCGAATGTCTGCCAAGACAGCACCCTGGGCTCTAGTTGGCGTCTCTTCAT 484  
QY 242 CAAGACGGAACCTTCCAGCCCAAGCTCCCTGACGAGCAGCGTCAACACACAGACCCCTGG 301  
Db 485 CAAGACGGAACCTTCCAGCCCAAGCTCCCTGACGAGCAGCGTCAACACACAGACCCCTGG 544  
QY 302 TGGCTCTTCAGACGCCAGTGGAGACTACAGTTCAACCATGAATGGCCATCAGACGAGACT 361  
Db 545 CGGCTCTGCGAGCGCAGCGAGTGTGCTTTGGCATTTGCCCTGACCAACCAACCGGCTCT 604  
QY 362 TGACTCGCCACCTCTTACCTTTCTCTCTATCTCTGAGGAGTGTGGGCTGTCAAGAA 421  
Db 605 GGACTCGCCGCTATGTTGCGAGGTGC-----GGGCTGGAGGCAACCGGCTGCGCAA 658  
QY 422 ACTGTATGATGACTGCTCCAGACCATTTGTGAAGATCCCAACCAAGTGTGAATACAT 481  
Db 659 GAGCTACGAGACTGTACTAGTGTATCATGAGAGACTCCGCATCAAAATGCGAGTACAT 718  
QY 482 GCTCAACTCGATGCCCAAGAGACTGTGTTAGTGTGTGTGATCGCTTCTGGTACCA 541  
Db 719 GCTTAACCCCATCCCAAGCGCCTGTGCTCTGTGCGGGACATGCTCTGCTACCA 778  
QY 542 CTATGGGCTAGCATCATGTGAGAGCTGCAAGGCATTTCTCAAGAGCAATTCAGAGCAA 601  
Db 779 CTACGAGTGGCCCTCTCGAGGCTTGCAGAGCGCTTCTTCAAGAACCATTCAGAGCAA 838  
QY 602 TATGAATACAGCTGCCCTGCCCAAGATGAATGTGAATCAAAAGCGCAGACGTAATC 661  
Db 839 CATCGATCAACTGCCCCGCCCAACCAATGAATGTGAATCAAAAGCGCAGAGTGC 898  
QY 662 CTGCCAGGCTTGGCGCTTCAATGAATGTGTTAAAGTGGGCATGCTGAAGAGGGTGGC 721  
Db 899 CTGTACGCTGTGCCGATTCATGAATGCTCAAAAGTGGGATGTGAAGGAGGTGTGC 958  
QY 722 TCTTGAAGAGTACGTGAGGTGGGAGAGTACAAAGCGCAGAGTAGATGCGAGAACAG 781  
Db 959 CCTGACCGAGTTCGAGAGGCGCGCAGAAGTACAAAGCGAGGCTGAGTTCGAGAACAG 1018  
QY 782 CCCATACCTGAACCCCTCAGCTGTTCAAGCCAGCCAAAAGCCATTAACAAGATTGCTC 841  
Db 1019 CCCCTACCTGAACCTGCGGATTTCCCACTGCTAAAGCCATGAAGATGCTCTC 1078  
QY 842 ACATTTGTTGTGCTGTAACCGGAGAGATCTATGCCATGCTGACCCCTACTGTCCCGCA 901  
Db 1079 GAATCTACTAGGGGTGAGCAGAGCAAGCTGTATGCTATGCTCCCAAGATATCCCGCA 1138  
QY 902 CAGTGAATCAAAAGCCCTCACTACCTGTGTGACTGTGGCCGACGAGAGTGTGTAT 961



Db 1139 GGGAGATATCAAGGCCCTGACCACTCTCTGTGAATTGGCAGATCGGAGCCTTGTCTCT 1198  
QY 962 CATTTGATGGCGGAAGCATATTCAGGCTTCTCCACGCTGTCCCTGGCGGACCAAGATGAG 1021  
Db 1199 CATCAACTGGGCCAAGACACATCCAGGCTTCCCAAGTGTGACACTGGGGACCAAGATGAG 1258  
QY 1022 CCTTCTGAGAGTGTCTTGGATGGAATTTGATCCCTTGTGTGTAACCGGTCTCTTTC 1081  
Db 1259 CCTGTCTGAGAGTGTCTTGGATGGAATTTGATCCCTTGTGTGTAACCGGTCTCTTTC 1318  
QY 1082 ATTTGAGATGAACCTTGTCTATGACAGATATATATATGACGAGACCAAGTCCCAATT 1141  
Db 1319 ATACGATGACAAAGCTGGCATACGCCGAGACTATATCATGATGAGGAACACTCTCGCCT 1378  
QY 1142 AGCAGGCTTCTTGTATCTAATATGCTATCTCTGACGCTGTAAAGAAATACAGAGCAT 1201  
Db 1379 GGTAGGGCTGTCTGACCTTTACCGGACATCTCTGACGCTGTGCGGAGGTACAGAAACT 1438  
QY 1202 GAAGCTGAGAAAAAGAAATTTGTACCCCTCAAGCTATAGCTTTGCTAATTCAGACTC 1261  
Db 1439 CAAGGTAGAGAAAGAAAGATTATGATCTCAAGGCCCTGCGCCCAACTCAGATTTC 1498  
QY 1262 CATGCATAGAAAGATGTTGAAGCCGTTCAAGAGCTTCAGATGCTTACATGAAGCGCT 1321  
Db 1499 GATGTACATGAGAACTGTGAGCGGTGAGAAAGCTCCAGACCTGCTGACGAGCGCT 1558  
QY 1322 GCAGGATATGAAGCTGGCCAGACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGAT 1381  
Db 1559 GCAGGATATGAGCTGAGTACGCGCCAGAGAGCGCGGAGGCGGCAAGCTGCTGCT 1618  
QY 1382 GACACTGCCACTCTCTGAGGACAGACTTACCAAGCGCTGACAGATTCTTACAACATCAA 1441  
Db 1619 GACGCTGCCCCCTGTGAGGACAGACAGCCGCCAAGCCGTGCAACATCTTACAGTGTGA 1678  
QY 1442 ACTAGAGGCAAGTCCCAATGCAAACTTTTGTGAAATGTTGGAAGGCCAAGGCTCTG 1501  
Db 1679 ACTGCGGGGCAAGGTGCCCATGCAAACTCTTCTGAGATGCTGGAAGGCCAAGGCTGTG 1738  
QY 1502 A 1502  
Db 1739 A 1739

RESULT 13  
B60472 704 bp DNA linear GSS 21-JUN-1998  
LOCUS CIT-HSP-2008K21.TRB CIT-HSP Homo sapiens genomic clone 2008K21,  
DEFINITION Genomic survey sequence.  
ACCESSION B60472  
VERSION B60472.1 GI:2615190  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 704)  
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
Simon,M. and Venter,J.C.  
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map  
BUILDING Building  
JOURNAL Unpublished (1997)  
COMMENT Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/cdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13 Reverse

FEATURES Class: BAC ends.  
source location/Qualifiers  
1..704  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7040519"  
/db\_xref="taxon:9606"  
/clone="2008K21"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_1ib="CIT-HSP"  
/note="Vector: pBelOBAC11; site\_1: HindIII; site\_2:  
HindIII"

Query Match 22.7%; Score 679; DB 28; Length 704;  
Best Local Similarity 99.6%; Pred. No. 2.8e-130;  
Matches 702; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 1293 AAGCTTCAGGATGCTTACATGAGAGGCTGACAGGATATGAGCTGGCCAGACATGGAA 1352  
Db 2 AAGCTTCAGGATGCTTACATGAGAGGCTGACAGGATATGAGCTGGCCAGACATGGAA 61  
QY 1353 GACCTCGTGCAGCTGGCAAGATGCTGATGACACTGCCACTCTGAGGCGACACTTACC 1412  
Db 62 GACCTCGTGCAGCTGGCAAGATGCTGATGACACTGCCACTCTGAGGCGACACTTACC 121  
QY 1413 AAGCCGTGCAGCATTTCTACACATCAAACTAGAAAGGCAAGTCCCAATGACAAACTT 1472  
Db 122 AAGCCGTGCAGCATTTCTACACATCAAACTAGAAAGGCAAGTCCCAATGACAAACTT 181  
QY 1473 TTTTGGAAATGTTGAGGCCAAGCTCTGAATAAAGCTCCCTGGGCTTCCCATCTTC 1532  
Db 182 TTTTGGAAATGTTGAGGCCAAGCTCTGAATAAAGCTCCCTGGGCTTCCCATCTTC 241  
QY 1533 ATGTTGAAAAAGGAAAAATAAACCAAGAGTGATGTGCAAGAACTTAGAGTTAGTTAA 1592  
Db 242 ATGTTGAAAAAGGAAAAATAAACCAAGAGTGATGTGCAAGAACTTAGAGTTAGTTAA 301  
QY 1593 CAACATCAAAAATCAACAGACTGCACTGATTAATTAGCAGCAAGACTATGAAGCAGCTTT 1652  
Db 302 CAACATCAAAAATCAACAGACTGCACTGATTAATTAGCAGCAAGACTATGAAGCAGCTTT 361  
QY 1653 CAGATTCCTCCATAGGTTCTGATGAG-TTCTTTCTACTTTCTCATCATCTTCTTCT 1711  
Db 362 CAGATTCCTCCATAGGTTCTGATGAGTTCTTTCTACTTTCTCATCATCTTCTTCT 421  
QY 1712 CTTTCTCCCAACATTTCTCTTCTCTTTAATTTTCTCTCTTCTCTTCTTCTTCTTCT 1771  
Db 422 CTTTCTCCCAACATTTCTCTTCTCTTTAATTTTCTCTCTTCTCTTCTTCTTCTTCT 481  
QY 1772 AATTCTTGTCTTCTTCAATCTAGTTCCCATTTCTCTTAAATTTCTTCCCGTCTGCGTG 1831  
Db 482 AATTCTTGTCTTCTTCAATCTAGTTCCCATTTCTCTTAAATTTCTTCCCGTCTGCGTG 541  
QY 1832 CCTTCTTCTTCTTCTTACCTACTCATCTCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1891  
Db 542 CCTTCTTCTTCTTCTTACCTACTCATCTCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 601  
QY 1892 TAAATTTGAATAGCTTTAGTTTAAAAAATATCCCTCCCTTCCCTTCTTCTTCTTCT 1951  
Db 602 TAAATTTGAATAGCTTTAGTTT-AAAAAATATCCCTCCCTTCTTCTTCTTCTTCTTCT 659  
QY 1952 TTCTTCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1996  
Db 660 TTCTTCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 704

RESULT 14  
AY409736 1296 bp DNA linear GSS 16-DEC-2003  
LOCUS AY409736  
DEFINITION Mus musculus ESRRB gene, VIRUTAL TRANSCRIPT, partial sequence,  
genomic survey sequence.

ACCESSION AY409736  
VERSION AY409736.1 GI:39765704  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS 1 (bases 1 to 1296)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003).  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 1296)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..1296  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>1296  
/gene="ESRRB"  
/locus\_tag="HCM3656"  
ORIGIN  
Query Match 22.5%; Score 670.4; DB 29; Length 1296;  
Best Local Similarity 70.3%; Pred. No. 1.6e-128;  
Matches 915; Conservative 0; Mismatches 381; Indels 6; Gaps 1;  
QY 195 ATGTCAACAAGATCGACACATTCATCCAGCTGTCGTCCTTCATCAAGACGGAACCT 254  
Db 1 ATGTCTCGAAGACAGGACCTGGGCTTAGTTGGGCTCCTTCATCAAGACGGAACCA 60  
QY 255 TCCAGCCAGCCTCCCTGACGACAGCTGCAACCAACAGCCCTGGTGGCTCTTCAGAC 314  
Db 61 TCCAGCCCTCTCGGCGATGAGCCCTCAGCCACACAGCCCAAGCGGCTGTCGGAC 120  
QY 315 GCCAGTGGAGCTACAGTTCAACCATGATGCCATCAGAACGAGCTGACTGCCACCT 374  
Db 121 GCCAGTGGTGGCTTTGGCATTTGCCCTGACCAACGCAACGGTGTGGACTGCCGCCCT 180  
QY 375 CTCTACCTCTGCTCCTATCCTGGAGGTAGTGGCCCTGTCAAGAACTGTATGATGAC 434  
Db 181 ATGTTCCGAGGTGC-----GGGGCTGGAGGCCAACCCGTGCCGCAAGACTACGAGAC 234  
QY 435 TGCTCCAGCACCATTTGTAAGATCCCCAGCAAGTGTGATATACATGCTCAACTCGATG 494  
Db 235 TGTACTAGTGTATCATGAGGAGCTCCGCCATCAAAATGCGATACATGCTTAAAGCCCATC 294  
QY 495 CCCAAGAGACTGTGTTTGTGTGTGTGACATCGCTTCTGGGTACCACTATGGGGTAGCA 554  
Db 295 CCCAAGCGCCTGTGCTGTGTGTGGGAGCAATTGCTGTGCTACCACTACGGAAGTGGCC 354  
QY 555 TCATGTGAAGCCTGCAAGGCATTCTTCAAGAGCAATTCAGGCAATATAGAAATACAGC 614  
Db 355 TCCTGCGAGGCTTCAAGCGCTTCTTCAAGGAACCATTCAGGCAACATCGATCAAC 414  
QY 615 TGCCCTGCCAAGATGAATGTGAATCACAAGGCGCAGACGTAAATCCTGCCAGGCTTGC 674  
Db 415 TGCCCGGCCACCAATGATGTGAGATCACCAACGAGAGCGCAAGTCTGTCAAGCCTGC 474  
QY 675 CGCTTCATGAAGTGTAAAGTGGGCGATGCTGAAGAAGGGGTGCGTCTTGACAGAGTA 734

Db 475 CGATTCATGAATGCTCAAGTGGGAGTGTGAAGAGTGTGCGCTTGACCGAGTT 534  
QY 735 CGTGAGGTCCGCGAAGATACAAAGCCGACGATAGTCGGAGAACAGCCCATACCTGAAC 794  
Db 535 CGAGGAGGCCCGCAGAAATACAAAGCGACGGCTGGAATTCGGAAGAACGCCCTACCTGAAC 594  
QY 795 CCTCAGCTGTTTCAGCCAGCCAAAAGCCATATACAGATTTCTCACAATTTGTGGTG 854  
Db 595 CTGCCGATTTCCCACTGCTAAAGGCCATTTGACTAAGATCGTCTCGAATCTACTAGGG 654  
QY 855 GCTGAACCGGAGAAATCTATGCCATGCTGACCTGACCTGATCCCGCAGACGTGACATCAA 914  
Db 655 GTTGAAGCAGACAAGCTGATGCTATGCTGCTCCCAAGATATCCCCGAGGAGATATCAAG 714  
QY 915 GCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGGTGTATCATTTGATGGCG 974  
Db 715 GCCCTGACCACTCTCTGTAATGGCAGATCGGAGCTGTGTCTCCTCATCAACTGGGCC 774  
QY 975 AAGCATATTCAGGCTTCTCCACGCTGTCCCTGGCGACCATGAGCCTTCTGCAGAGT 1034  
Db 775 AAGCATATCCAGGCTTCCCACTGTGACACTTGGGGACCATGAGCCTGTGCAGAGT 834  
QY 1035 GCTTGGATGAAATTTGATCTTGTGCTGTATACCGGCTCTTTCATTTGAGGATGAA 1094  
Db 835 GCTGATGAGATTTCTATCTTGGGCATCGTGTACCGCTGCTCCCATATGATGACAAG 894  
QY 1095 CTGTCTATGACAGACGATTAATTAATGACGAGAACACAGTCCAAATTAGCAGGCTTCT 1154  
Db 895 CTGCATATCCCGAGACTATATCATGATGAGGAACACTCTCGCTGTAGGCGTCTGT 954  
QY 1155 GATCTAAATATGCTATCTGTCAGCTGTGTAAGAAATACAGACATGAAAGCTGGAATA 1214  
Db 955 GACCTTAACGAGCATCTCTGACCTGTGCGCAGGTACAGAAACTCAAGGTAGAGAAG 1014  
QY 1215 GAAGAAATTGTCACCCCTCAAGCTATAGCTCTTGCTAATTCAGACTCCATGACATAGAA 1274  
Db 1015 GAAGAGTTTATGATCTCAAGGCCCTGGCCCTCGCCAACTGATGATGATGAG 1074  
QY 1275 GATGTTGAAGCCGTTCAAGAGCTTCAAGATGTCTTACATGAAGCGCTGCAGGATTATGAA 1334  
Db 1075 AACTTGAGGCGGTGCAAGAGCTCCAGAACCTGCTGCAAGAGCGCTGCAGGACTATGAG 1134  
QY 1335 GTTGCCAGCACATGGAAGACCTGCTGAGCTGCGCAAGATGCTGATGACACTGCCACTC 1394  
Db 1135 CTGATCAGCGCCACGAGAGCCCGGAGGGCGGCAAGCTGCTGACGCTGCCCTG 1194  
QY 1395 CTGAGCAGACCTCTACCAAGCCGTCAGCACTTCTACAACTCAAACTAGAAAGGCAAA 1454  
Db 1195 CTGAGCAGACAGCCGCCAAAGCCGTGCAACTTCTACAGTGTGAAACTGCAGGGCAAG 1254  
QY 1455 GTCCCAATGCACAACTTTTGTGAAATGTTGGAGGCCAAG 1496  
Db 1255 GTGCCCATGCACAACTCTTCTGAGATGCTGAGGCCAAAG 1296  
RESULT 15  
CB989379 780 bp mRNA linear EST 01-MAY-2003  
LOCUS CB989379  
DEFINITION AGENCOURT 13902955 NIH\_MGC\_147 Homo sapiens cDNA clone  
IMAGE:30345143 5', mRNA sequence.  
ACCESSION CB989379  
VERSION CB989379.1 GI:30283899  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Stefan Hansson  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
and advice from Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: NDAM382 row: 1 column: 24  
High quality sequence stop: 606.

FEATURES  
source  
1. .780  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30345143"  
/tissue\_type="Human Placenta"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH MGC 147"  
/note="Organ: placenta; Vector: pBluescriptR; Site\_1:  
all-XhoI; Site 2: BamH; Oligo-dr primed using primer  
5'-TTTTTTTTTTT-3', size-selected for average  
insert size 2.3 kb and normalized to ROT 5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NHGRI/NHGRI, National Institutes of Health). Note: This is  
a NIH\_MGC library."

ORIGIN  
Query Match 22.4%; Score 667.6; DB 14; Length 780;  
Best Local Similarity 98.4%; Pred. No. 6.4e-128;  
Matches 685; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 17 GTGGAATTCGGCTTGTCTACTAGAGAACTTTGTTAATTGCACTGTGCTCTGCAAGG 76  
Db 85 GTGCCTATAGGCTTGTCTACTAGAGAACTTTGTTAATTGCACTGTGCTCTGCAAGG 144  
QY 77 AAACCTTGATTATAGCTGGGTCACAAATATAGTTGCCGTCACATGATTCGGT 136  
Db 145 AAACCTTGATTATAGCTGGGTCACAAATATAGTTGCCGTCACATGATTCGGT 204  
QY 137 AGAAGCTTGCCCTTCCCTGAATCTTTCCCTGCACCTAGAGAGAGCTTCTCGAGAAT 196  
Db 205 AGAAGCTTGCCCTTCCCTGAATCTTTCCCTGCACCTAGAGAGAGCTTCTCGAGAAT 264  
QY 197 GTCAACAAGATGACACATGATTCAGCTGTGCTCCTTCATCAAGACGGAACCTTC 256  
Db 265 GTCAACAAGATGACACATGATTCAGCTGTGCTCCTTCATCAAGACGGAACCTTC 324  
QY 257 CAGCCAGCCTCCCTGACGACAGGTCACACCAACACAGCCCTGGTGGCTCTTCAGACGC 316  
Db 325 CAGCCAGCCTCCCTGACGACAGGTCACACCAACACAGCCCTGGTGGCTCTTCAGACGC 384  
QY 317 CAGTGGAGGCTACAGTTCACACATGATGGCCATCAGAACGGAACCTTCGACCTCT 376  
Db 385 CAGTGGAGGCTACAGTTCACACATGATGGCCATCAGAACGGAACCTTCGACCTCT 444  
QY 377 CTACCTTCTGCTCTATCCTGGAGGTAGTGGCCCTGTCAGAAACTGTATGATGACTG 436  
Db 445 CTACCTTCTGCTCTATCCTGGAGGTAGTGGCCCTGTCAGAAACTGTATGATGACTG 504  
QY 437 CTCAGACCACTTGTGAAGATCCCAAGCAAGTGTGAATATATGCTCACTCGATGCC 496  
Db 505 CTCAGACCACTTGTGAAGATCCCAAGCAAGTGTGAATATATGCTCACTCGATGCC 564  
QY 497 CAAGAGACTGTGTTAGTGTGTGATGACATCGCTTCTGGGTACCACTATGGGGTAGCATC 556  
Db 565 CAAGAGACTGTGTTAGTGTGTGATGACATCGCTTCTGGGTACCACTATGGGGTAGCATC 624  
QY 557 ATGTGAAGCCTGCAAGGCAATCTTCAAGAGACCAATCAAGGCAATATAGAAATACAGCTG 616

Db 625 ATGTGAAGCCTGCAAGGCAATCTTCAAGAGACCAATCAAGGCAATATAGAAATACAGCTG 684  
QY 617 CCCTGCCAGCAATGAATGTGAATCACAAGCGCAGACGTAATCCTGCC--AGGCTTGC 674  
Db 685 CCCTGCCAGCAATGAATGTGAATCACAAGCGCAGACGTAATCCTGCCAGGCTTGC 744  
QY 675 CGCTTCATGAAGTCTTAAAGTGGGCAATGCTGAAA 710  
Db 745 CGCTTCATGAAGTCTTAAAGTGGGCAATGCTGAAA 780

Search completed: August 10, 2004, 04:55:27  
Job time : 7350 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2004, 23:15:03 ; Search time 208 Seconds  
(without alignments)  
7964.079 Million cell updates/sec

Title: US-10-054-841-3  
Perfect score: 2985  
Sequence: 1 gcggcgccgcagtggtgtg.....actggcgccgcgcttgat 2985

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1559	52.2	1615	3	US-09-040-508-1	Sequence 1, Appli
2	1559	52.2	1615	4	US-09-500-654-1	Sequence 1, Appli
3	443.2	14.8	536	3	US-09-040-508-3	Sequence 3, Appli
4	443.2	14.8	536	4	US-09-500-654-3	Sequence 3, Appli
5	421.4	14.1	2539	4	US-09-620-312D-454	Sequence 454, App
6	162.6	5.4	7218	1	US-08-232-463-14	Sequence 14, Appli
7	136.6	4.6	4963	1	US-08-076-726-16	Sequence 16, Appli
8	136.6	4.6	4963	1	US-08-260-452-9	Sequence 9, Appli
9	136.6	4.6	4963	2	US-08-481-970-9	Sequence 9, Appli
10	136.6	4.6	4963	2	US-08-897-719-9	Sequence 9, Appli
11	136.6	4.6	4963	3	US-09-163-269-9	Sequence 9, Appli
12	136.6	4.6	6450	3	US-09-041-886-34	Sequence 34, Appli
13	136.6	4.6	6450	4	US-08-453-998-1	Sequence 1, Appli
14	136.6	4.6	6450	4	US-10-027-983-3	Sequence 3, Appli
15	131.6	4.4	472	4	US-09-833-381-269	Sequence 269, App
16	131.6	4.4	472	4	US-09-833-381-270	Sequence 270, App
17	125.8	4.2	1757	4	US-08-216-592A-3	Sequence 3, Appli
18	125.8	4.2	2130	1	US-07-952-800-1	Sequence 1, Appli
19	125.8	4.2	2204	1	US-07-952-800-3	Sequence 3, Appli
20	125.8	4.2	2219	1	US-08-336-408B-7	Sequence 7, Appli
21	125.8	4.2	2219	5	PCT-US91-00399-7	Sequence 7, Appli
22	124.6	4.2	2928	2	US-08-095-728B-3	Sequence 3, Appli
23	124.6	4.2	2928	5	PCT-US92-02320A-3	Sequence 3, Appli
24	124.6	4.2	2940	2	US-08-592-383-1	Sequence 1, Appli
25	124.6	4.2	3036	1	US-08-306-691B-52	Sequence 52, Appli
26	124.6	4.2	3036	2	US-08-095-728B-1	Sequence 1, Appli
27	124.6	4.2	3036	5	PCT-US92-02320A-1	Sequence 1, Appli

28	124.6	4.2	3511	3	US-08-892-747-13	Sequence 13, Appli
29	124.2	4.2	2285	4	US-08-216-592A-1	Sequence 1, Appli
30	123.6	4.1	1576	6	5260432-1	Patent No. 5260432
31	123	4.1	2940	6	5171671-1	Patent No. 5171671
32	122.6	4.1	1594	4	US-09-016-434-1059	Sequence 1059, Ap
33	120.4	4.0	1883	4	US-08-216-592A-5	Sequence 5, Appli
34	120.4	4.0	2095	1	US-08-336-408B-3	Sequence 3, Appli
35	120.4	4.0	2095	5	PCT-US91-00399-3	Sequence 3, Appli
36	118.2	4.0	1866	1	US-08-336-408B-1	Sequence 1, Appli
37	118.2	4.0	1866	5	PCT-US91-00399-1	Sequence 1, Appli
38	113.8	3.8	1251	4	US-09-608-088-2	Sequence 2, Appli
39	113.8	3.8	1257	4	US-09-608-088-20	Sequence 20, Appli
40	113.8	3.8	1434	4	US-09-608-088-1	Sequence 1, Appli
41	113.8	3.8	1460	2	US-08-836-620A-4	Sequence 4, Appli
42	113.8	3.8	1647	3	US-09-139-617-2	Sequence 2, Appli
43	113.8	3.8	1647	4	US-09-561-741A-2	Sequence 2, Appli
44	113.8	3.8	1647	4	US-09-558-795-2	Sequence 2, Appli
45	113.8	3.8	1898	4	US-09-608-088-24	Sequence 24, Appli

ALIGNMENTS

RESULT 1  
US-09-040-508-1  
: Sequence 1, Application US/09040508  
: Patent No. 6069239  
: GENERAL INFORMATION:  
: APPLICANT: Mathias, Stephen L.  
: TITLE OF INVENTION: NOVEL COMPOUNDS  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: RATNER & PRESTIA  
: STREET: P.O. BOX 980  
: CITY: VALLEY FORGE  
: STATE: PA  
: COUNTRY: USA  
: ZIP: 19482  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FastSeq for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/040,508  
: FILING DATE: 17-MAR-1998  
: CLASSIFICATION:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: GB 9705451.4  
: FILING DATE: 17-MAR-1997  
: APPLICATION NUMBER: GB 9803289.9  
: FILING DATE: 16-FEB-1998  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PRESTIA, PAUL F  
: REGISTRATION NUMBER: 23,031  
: REFERENCE/DOCKET NUMBER: GH-30008  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 610-407-0700  
: TELEFAX: 610-407-0701  
: TELEX: 846169  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1615 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA  
: US-09-040-508-1

Query Match 52.2%; Score 1559; DB 3; Length 1615;  
Best local Similarity 99.7%; Pred. No. 0;  
Matches 1562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 17 GTGAATTGGCTGTGCTAGAGAAACATTTGTATTGACCTGTCTGTCAAGG 76  
Db 49 GTGCCTATAGGCTGTGCTAGAGAAACATTTGTATTGACCTGTCTGTCAAGG 108  
QY 77 AAACCTTGATTATAGCTGGGGTGACAAATAATAGTTGCCGTGCACATGGAATTCGGT 136  
Db 109 AAACCTTGATTATAGCTGGGGTGACAAATAATAGTTGCCGTGCACATGGAATTCGGT 168  
QY 137 AGAAGCTTGGCTTCTGTAATCTTTTCCCTGCACTACGAGGAAGAGCTTCTGTGAGAAT 196  
Db 169 AGAAGCTTGGCTTCTGTAATCTTTTCCCTGCACTACGAGGAAGAGCTTCTGTGAGAAT 228  
QY 197 GTCAAAACAAGATCGACACATTTGATTCAGCTGTTCCTTCTCATCAAGACGGAACCTTC 256  
Db 229 GTCAAAACAAGATCGACACATTTGATTCAGCTGTTCCTTCTCATCAAGACGGAACCTTC 288  
QY 257 CAGCCCGAGCTCCCTGACGGACAGCGTCAACCAACAGCCCTGGTGGCTCTCAGACGCG 316  
Db 289 CAGCCCGAGCTCCCTGACGGACAGCGTCAACCAACAGCCCTGGTGGCTCTCAGACGCG 348  
QY 317 CAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGACTTGACTCGCCACCTCT 376  
Db 349 CAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGACTTGACTCGCCACCTCT 408  
QY 377 CTACCTCTTCTGCTCCTATCCCTGGAGTAGTGGCCCTGTCAAGAACTGTATGATGACTG 436  
Db 409 CTACCTCTTCTGCTCCTATCCCTGGAGTAGTGGCCCTGTCAAGAACTGTATGATGACTG 468  
QY 437 CTCCAGCAACCATTTGTTGAAGATCCCGACAGCAAGTGTGAATACATGCTCACTCGATGCC 496  
Db 469 CTCCAGCAACCATTTGTTGAAGATCCCGACAGCAAGTGTGAATACATGCTCACTCGATGCC 528  
QY 497 CAAGAGACTGTGTTTGTAGTGTGTGTGACATCGCTTCTGGGTACCACTATGGGGTAGCATC 556  
Db 529 CAAGAGACTGTGTTTGTAGTGTGTGTGACATCGCTTCTGGGTACCACTATGGGGTAGCATC 588  
QY 557 ATGTGAAGCCTGCAAGGCATTCTTCAAGAGACAAATTCAGGCAATATAGAAATACAGCTG 616  
Db 589 ATGTGAAGCCTGCAAGGCATTCTTCAAGAGACAAATTCAGGCAATATAGAAATACAGCTG 648  
QY 617 CCCTGCCACGAATGAATGTGAATCACAAAGCGCAGACGTAATCTCCAGGCTTGCCG 676  
Db 649 CCCTGCCACGAATGAATGTGAATCACAAAGCGCAGACGTAATCTCCAGGCTTGCCG 708  
QY 677 CTTCATGAAGTGTAAAAAGTGGGCATGCTGAAGAAGGGGTGCGTCTTGACAGAGTACG 736  
Db 709 CTTCATGAAGTGTAAAAAGTGGGCATGCTGAAGAAGGGGTGCGTCTTGACAGAGTACG 768  
QY 737 TGGAGGTGCGCAGAAGTACAAAGCGCAGATAGATGCGGAGAACAGCCCATACCTGAACC 796  
Db 769 TGGAGGTGCGCAGAAGTACAAAGCGCAGATAGATGCGGAGAACAGCCCATACCTGAACC 828  
QY 797 TCAGCTGGTTCAGCCAGCCAAAAGCCATATATACAGATTTCTCACATTTGTTGGTGGC 856  
Db 829 TCAGCTGGTTCAGCCAGCCAAAAGCCATATATACAGATTTCTCACATTTGTTGGTGGC 888  
QY 857 TGAACCGGAGAAGATCTATGCCATGCCCTGACCCCTACTGTCCCGACAGTGACATCAAAAGC 916  
Db 889 TGAACCGGAGAAGATCTATGCCATGCCCTGACCCCTACTGTCCCGACAGTGACATCAAAAGC 948  
QY 917 CCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGGTTATCATTTGATGGCGAA 976  
Db 949 CCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGGTTATCATTTGATGGCGAA 1008  
QY 977 GCATATTCAGGCTTCTCAGCGTGTCCCTGGCGGACAGATGAGCCTTCTGCAAGTGC 1036  
Db 1009 GCATATTCAGGCTTCTCAGCGTGTCCCTGGCGGACAGATGAGCCTTCTGCAAGTGC 1068  
QY 1037 TTGGATGGAATTTTGAATCTTGGTGTGCTATACCGGTCTCTTCAATTTGAGATGAAT 1096  
Db 1069 TTGGATGGAATTTTGAATCTTGGTGTGCTATACCGGTCTCTTCAATTTGAGATGAAT 1128  
QY 1097 TGTCTATGACAGCATATATATATGACGAGAACCAAGTCCAAATTAGCAGGCTTCTTGA 1156

Db 1129 TGTCTATGACAGCATATATATATGACGAGAACCAAGTCCAAATTAGCAGGCTTCTTGA 1188  
QY 1157 TCTAAATAATGCTATCCCTGACAGCTGGTAAAGAAATACAAAGACATGAAGCTGGAATAAG 1216  
Db 1189 TCTAAATAATGCTATCCCTGACAGCTGGTAAAGAAATACAAAGACATGAAGCTGGAATAAG 1248  
QY 1217 AGAATTTGTCAACCCCTCAAAGCTATAGCTCTGTCTAATTCAGACTCCATGACATAGAAGA 1276  
Db 1249 AGAATTTGTCAACCCCTCAAAGCTATAGCTCTGTCTAATTCAGACTCCATGACATAGAAGA 1308  
QY 1277 TGTGAAGCCGTTCAAGAGCTTCAGGATGTCTTACATGAAGCGCTGCAAGATTATGAAGC 1336  
Db 1309 TGTGAAGCCGTTCAAGAGCTTCAGGATGTCTTACATGAAGCGCTGCAAGATTATGAAGC 1368  
QY 1337 TGGCCAGCACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGATGACACTGCCACTCCT 1396  
Db 1369 TGGCCAGCACATGGAAGACCCCTGCTGAGCTGGCAAGATGCTGATGACACTGCCACTCCT 1428  
QY 1397 GAGGACAGCTCTACCAAGGCGTGACGATTTCTACACATCAAACTAGAAGGCAAAAGT 1456  
Db 1429 GAGGACAGCTCTACCAAGGCGTGACGATTTCTACACATCAAACTAGAAGGCAAAAGT 1488  
QY 1457 CCCAATGACAAACTTTTTTTGGAATGTTGAGGCGCAAGGTGACTAAAAAGCTCCCTG 1516  
Db 1489 CCCAATGACAAACTTTTTTTGGAATGTTGAGGCGCAAGGTGACTAAAAAGCTCCCTG 1548  
QY 1517 GGCTTCCCATCCTTCATGTTGAAAAAGGAAAAATAAACCCAGAGTGAATGTGGAAGAA 1576  
Db 1549 GGCTTCCCATCCTTCATGTTGAAAAAGGAAAAATAAACCCAGAGTGAATGTGGAAGAA 1608  
QY 1577 CTTAGAG 1583  
Db 1609 CTTAGAG 1615

RESULT 2  
US-09-500-654-1  
: Sequence 1, Application US/09500654  
: Patent No. 6359116  
: GENERAL INFORMATION:  
: APPLICANT: Stephen L. Mathias  
: TITLE OF INVENTION: No. 6359116e1 Compounds  
: FILE REFERENCE: GP-30008-D1  
: CURRENT APPLICATION NUMBER: US/09/500,654  
: EARLIER FILING DATE: 2000-02-09  
: EARLIER APPLICATION NUMBER: 9705451.4  
: EARLIER FILING DATE: 1997-03-17  
: EARLIER APPLICATION NUMBER: 9803289.9  
: NUMBER OF SEQ ID NOS: 4  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 1  
: LENGTH: 1615  
: TYPE: DNA  
: ORGANISM: HOMO SAPIENS  
: US-09-500-654-1  
Query Match 52.2%; Score 1559; DB 4; length 1615;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 17 GTGAATTGGCTGTGCTAGAGAAACATTTGTATTGACCTGTCTGTCAAGG 76  
Db 49 GTGCCTATAGGCTGTGCTAGAGAAACATTTGTATTGACCTGTCTGTCAAGG 108  
QY 77 AAACCTTGATTATAGCTGGGGTGACAAATAATAGTTGCCGTGCACATGGAATTCGGT 136  
Db 109 AAACCTTGATTATAGCTGGGGTGACAAATAATAGTTGCCGTGCACATGGAATTCGGT 168  
QY 137 AGAAGCTTGGCTTCTGTAATCTTTTCCCTGCACTACGAGGAAGAGCTTCTGTGAGAAT 196  
Db 169 AGAAGCTTGGCTTCTGTAATCTTTTCCCTGCACTACGAGGAAGAGCTTCTGTGAGAAT 228

QY 197 GTCAACAAGATCGACATTTGATTCAGCTGTTCCTTCATCAAGACGGAACCTTC 256  
DB 229 GTCAACAAGATCGACATTTGATTCAGCTGTTCCTTCATCAAGACGGAACCTTC 288  
QY 257 CAGCCAGCCTCCCTGACGACAGCGTCAACCAACCAAGCCTGTGGCTCTTGACGCG 316  
DB 289 CAGCCAGCCTCCCTGACGACAGCGTCAACCAACCAAGCCTGTGGCTCTTGACGCG 348  
QY 317 CAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGAATTGACTCGCCACTCT 376  
DB 349 CAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGAATTGACTCGCCACTCT 408  
QY 377 CTACCTTCTGCTCTATCTGGGAGGTAGTGGCCCTGTGAGGAACTGTATGATGACTG 436  
DB 409 CTACCTTCTGCTCTATCTGGGAGGTAGTGGCCCTGTGAGGAACTGTATGATGACTG 468  
QY 437 CTCAGACCACTTGTGTAAGATCCCGACCAAGTGTGAATACATGCTCAACTCGATGCC 496  
DB 469 CTCAGACCACTTGTGTAAGATCCCGACCAAGTGTGAATACATGCTCAACTCGATGCC 528  
QY 497 CAAGAGCTGTGTTAGTGTGTGATCATCGCTTCTGGGTACCACTATGGGTAGCATC 556  
DB 529 CAAGAGCTGTGTTAGTGTGTGATCATCGCTTCTGGGTACCACTATGGGTAGCATC 588  
QY 557 ATGTGAAGCCTGCAAGGCATTTCTCAAGAGACATTCAGGCAATATAGATACAGCTG 616  
DB 589 ATGTGAAGCCTGCAAGGCATTTCTCAAGAGACATTCAGGCAATATAGATACAGCTG 648  
QY 617 CCTGCGCACGAATGATGTGAATTCACAAAGCGCAGACGTAAATCTGCGAGGCTTGCGG 676  
DB 649 CCTGCGCACGAATGATGTGAATTCACAAAGCGCAGACGTAAATCTGCGAGGCTTGCGG 708  
QY 677 CTTCAATGAAGTGTTTAAAGTGGGCATGCTGAAAGAAAGGGGTGCTTGTACAGAGTACG 736  
DB 709 CTTCAATGAAGTGTTTAAAGTGGGCATGCTGAAAGAAAGGGGTGCTTGTACAGAGTACG 768  
QY 737 TGGAGTGGGACGAAGTACAAAGCGCAGAGTAGATGCGGAGAACCCCATACCTGAACCC 796  
DB 769 TGGAGTGGGACGAAGTACAAAGCGCAGAGTAGATGCGGAGAACCCCATACCTGAACCC 828  
QY 797 TCAGCTGTTTCAAGCCAAAGCCAAATATACAAAGATTTCTCAATTTGTTGTTGCG 856  
DB 829 TCAGCTGTTTCAAGCCAAAGCCAAATATACAAAGATTTCTCAATTTGTTGTTGCG 888  
QY 857 TGAACCGGAGAAAGATCTATGCCATGCTGACCCCTACTGTCCCGACAGTACATCAAGC 916  
DB 889 TGAACCGGAGAAAGATCTATGCCATGCTGACCCCTACTGTCCCGACAGTACATCAAGC 948  
QY 917 CCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGTGTATCATTTGATGGCGAA 976  
DB 949 CCTCACTACACTGTGTGACTTGGCCGACCGAGAGTGTGTGTATCATTTGATGGCGAA 1008  
QY 977 GCATATTCAGGCTTCTCCACGCTGTCCCTGGCGGACAGATGAGCCTTCTGCAGAGTGC 1036  
DB 1009 GCATATTCAGGCTTCTCCACGCTGTCCCTGGCGGACAGATGAGCCTTCTGCAGAGTGC 1068  
QY 1037 TTGATGGAATTTTGTATCCTGTGTGTGTATACCGGTCTCTTCAATTTGAGATGAAC 1096  
DB 1069 TTGATGGAATTTTGTATCCTGTGTGTGTATACCGGTCTCTTCAATTTGAGATGAAC 1128  
QY 1097 TGTCTATGACAGCATTATATATATGACGAGACCAAGTCCAAATTAGAGGCTTCTTGA 1156  
DB 1129 TGTCTATGACAGCATTATATATATGACGAGACCAAGTCCAAATTAGAGGCTTCTTGA 1188  
QY 1157 TCTAATATATGCTATCTCGAGCTGTAAAGAAATACAAAGACATGAAGTGAAGAAAGA 1216  
DB 1189 TCTAATATATGCTATCTCGAGCTGTAAAGAAATACAAAGACATGAAGTGAAGAAAGA 1248  
QY 1217 AGAATTTGTACACCTCAAGCTATAGCTCTTGTCTAATTCAGACTCCATGCACTAGAAAGA 1276  
DB 1249 AGAATTTGTACACCTCAAGCTATAGCTCTTGTCTAATTCAGACTCCATGCACTAGAAAGA 1308

QY 1277 TGTGAAGCCGTTGAGAAAGCTTCAGGATGTCTTACATGAAGCGCTGCAGGATTAAGAC 1336  
DB 1309 TGTGAAGCCGTTGAGAAAGCTTCAGGATGTCTTACATGAAGCGCTGCAGGATTAAGAC 1368  
QY 1337 TGGCCAGACATGGAAGACCCCTCGTCGAGCTGGCAAGATGCTGATGACACTGCCACTCT 1396  
DB 1369 TGGCCAGACATGGAAGACCCCTCGTCGAGCTGGCAAGATGCTGATGACACTGCCACTCT 1428  
QY 1397 GAGGACAGCCTTACCAAGCCGTCAGACATTTCTACACATCAAACTAGAAAGCAAGT 1456  
DB 1429 GAGGACAGCCTTACCAAGCCGTCAGACATTTCTACACATCAAACTAGAAAGCAAGT 1488  
QY 1457 CCCAATGCACAAACTTTTGTGAAATGTTGAGGCCAAGTCTGACTAAAGCTCCCTG 1516  
DB 1489 CCCAATGCACAAACTTTTGTGAAATGTTGAGGCCAAGTCTGACTAAAGCTCCCTG 1548  
QY 1517 GGCCCTCCCATCCTTCATGTTGAAATAAGGAAATTAACCAAGATGATGTGAAGAA 1576  
DB 1549 GGCCCTCCCATCCTTCATGTTGAAATAAGGAAATTAACCAAGATGATGTGAAGAA 1608  
QY 1577 CTTAGAG 1583  
DB 1609 CTTAGAG 1615

RESULT 3  
US-09-040-508-3  
: Sequence 3, Application us/09040508  
: Patent No. 6069239  
: GENERAL INFORMATION:  
: APPLICANT: Mathias, Stephen L.  
: TITLE OF INVENTION: NOVEL COMPOUNDS  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: RATNER & PRESTIA  
: STREET: P.O. BOX 980  
: CITY: VALLEY FORGE  
: STATE: PA  
: COUNTRY: USA  
: ZIP: 19482  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FastSeq for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/040,508  
: FILING DATE: 17-MAR-1998  
: CLASSIFICATION:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: GB 9705451.4  
: FILING DATE: 17-MAR-1997  
: APPLICATION NUMBER: GB 9803289.9  
: FILING DATE: 16-FEB-1998  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PRESTIA, PAUL F  
: REGISTRATION NUMBER: 23,031  
: REFERENCE/DOCKET NUMBER: GH-30008  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 610-407-0700  
: TELEFAX: 610-407-0701  
: TELEX: 846169  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 536 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA  
: US-09-040-508-3  
Query Match 14.8%; Score 443.2; DB 3; Length 536;  
Best Local Similarity 96.1%; Pred. No. 3.7e-111;



```
Matches 497; Conservative 0; Mismatches 14; Indels 6; Gaps 4;
QY 17 GTGGAATTCGGCTTGTCTCACTAGAGAACATTTGTTAATGCACTGTGCTCTGTCAGG 76
Db 20 GTGCCTATAGGCTGTGTCACTAGAGAACATTTGTTAATGCACTGTGCTCTGTCAGG 79
QY 77 AAACCTTGAATTTATAGCTGGGGTGACAAATAA-TGGTTGCCGGTGCACATGATTC-G 134
Db 80 NAACCTTGAATTTATAGCTGGGGTGACAAATAAATTGGTTGCCGGTGCACATGATTCGG 139
QY 135 GTAGAACCTTTCCTTCTGTAATCTTTTCCCTGCACTACGAGGAAGAG---CTTCTCTGC 191
Db 140 GAAGAACCTTACCTTCTGTAATCTTTTCCCTGCACTACGAGGAAGAGCTTTCTCTGC 199
QY 192 AGAATGT-CAAAACAAGATCGACACATTTCCAGCTGTTCCCTTCATCAAGACGGA 250
Db 200 AGAATGTCCAAACAAGATCGACACATTTCCAGCTGTTCCCTTCATCAAGACGGA 259
QY 251 ACCTTCCAGCCAGCCTCCCTGACGGACAGCGTCAACCAACAGCCCTGGTGCTCTTC 310
Db 260 ACCTTCCAGCCAGCCTCCCTGACGGACAGCGTCAACCAACAGCCCTGGTGCTCTTC 319
QY 311 AGACGCCAGTGGGAGCTACAGTTCAACCATGAAATGCCCATCAGAACGCACTTGACTCGCC 370
Db 320 AGACGCCAGTGGGAGCTACAGTTCAACCATGAAATGCCCATCAGAACGCACTTGACTCGCC 379
QY 371 ACCTCTCTACCCCTTCTGCTCCTATCCTGGGAGGTAGTGGCCCTGTGCAAGAACTGTATGA 430
Db 380 ACCTCTCTACCCCTTCTGCTCCTATCCTGGGAGGTAGTGGCCCTGTGCAAGAACTGTATGA 439
QY 431 TGACTGCTCCAGCACCATTTGTTGAAGATCCCCAGACCAAGTGTGAATACATGCTCAACTC 490
Db 440 TGACTGCTCCAGCACCATTTGTTGAAGATCCCCAGACCAAGTGTGAATACATGCTCAACTC 499
QY 491 GATGCCCAAGAGACTGTGTTAGTGTGTGTGACATC 527
Db 500 GATGCCCAAGAGACTGTGTTAGTGTGTGTGACATC 536
```

```
RESULT 4
US-09-500-654-3
; Sequence 3, Application US/09500654
; Patent No. 6359116
; GENERAL INFORMATION:
; APPLICANT: Stephen L. Mathias
; TITLE OF INVENTION: No. 6359116el Compounds
; FILE REFERENCE: GP-30008-D1
; CURRENT APPLICATION NUMBER: US/09/500,654
; EARLIER FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 9705451.4
; EARLIER FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: 9803289.9
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 536
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (80)
US-09-500-654-3
```

```
Query Match 14.8%; Score 443.2; DB 4; Length 536;
Best Local Similarity 96.1%; Pred. No. 3.7e-111;
Matches 497; Conservative 0; Mismatches 14; Indels 6; Gaps 4;
```

```
QY 17 GTGGAATTCGGCTTGTCTCACTAGAGAACATTTGTTAATGCACTGTGCTCTGTCAGG 76
Db 20 GTGCCTATAGGCTGTGTCACTAGAGAACATTTGTTAATGCACTGTGCTCTGTCAGG 79
QY 77 AAACCTTGAATTTATAGCTGGGGTGACAAATAA-TGGTTGCCGGTGCACATGATTC-G 134
```

```
Db 80 NAACCTTGAATTTATAGCTGGGGTGACAAATAAATTGGTTGCCGGTGCACATGATTCGG 139
QY 135 GTAGAACCTTTCCTTCTGTAATCTTTTCCCTGCACTACGAGGAAGAG---CTTCTCTGC 191
Db 140 GAAGAACCTTACCTTCTGTAATCTTTTCCCTGCACTACGAGGAAGAGCTTTCTCTGC 199
QY 192 AGAATGT-CAAAACAAGATCGACACATTTCCAGCTGTTCCCTTCATCAAGACGGA 250
Db 200 AGAATGTCCAAACAAGATCGACACATTTCCAGCTGTTCCCTTCATCAAGACGGA 259
QY 251 ACCTTCCAGCCAGCCTCCCTGACGGACAGCGTCAACCAACAGCCCTGGTGCTCTTC 310
Db 260 ACCTTCCAGCCAGCCTCCCTGACGGACAGCGTCAACCAACAGCCCTGGTGCTCTTC 319
QY 311 AGACGCCAGTGGGAGCTACAGTTCAACCATGAAATGCCCATCAGAACGCACTTGACTCGCC 370
Db 320 AGACGCCAGTGGGAGCTACAGTTCAACCATGAAATGCCCATCAGAACGCACTTGACTCGCC 379
QY 371 ACCTCTCTACCCCTTCTGCTCCTATCCTGGGAGGTAGTGGCCCTGTGCAAGAACTGTATGA 430
Db 380 ACCTCTCTACCCCTTCTGCTCCTATCCTGGGAGGTAGTGGCCCTGTGCAAGAACTGTATGA 439
QY 431 TGACTGCTCCAGCACCATTTGTTGAAGATCCCCAGACCAAGTGTGAATACATGCTCAACTC 490
Db 440 TGACTGCTCCAGCACCATTTGTTGAAGATCCCCAGACCAAGTGTGAATACATGCTCAACTC 499
QY 491 GATGCCCAAGAGACTGTGTTAGTGTGTGTGACATC 527
Db 500 GATGCCCAAGAGACTGTGTTAGTGTGTGTGACATC 536
```

```
RESULT 5
US-09-620-312D-454
; Sequence 454, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR APPLICATION NUMBER: 2000-07-19
; PRIOR FILING DATE: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 454
; LENGTH: 2539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)..(1755)
US-09-620-312D-454
```







APPLICATION NUMBER: 08/076,327  
ATTORNEY/AGENT INFORMATION:  
NAME: Giulio A. DeConti, Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: BBI-013CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4963 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Human cytomegalovirus  
IMMEDIATE SOURCE:  
CLONE: PUHD BGR4  
US-08-260-452-9

Query Match 4.6%; Score 136.6; DB 1; Length 4963;  
Best Local Similarity 52.3%; Pred. No. 7.8e-27;  
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

QY 507 TGTTTAGTGTGTGACATCGCTTCTGGGTACCATATAGGGGTAGCATCATGTGAAGC 566  
Db 1018 TGTGCAGTGTGCAATGACATAGCTTCAAGCTACCATATAGAGTCTGTCTGTGAGGC 1077  
QY 567 TGCAGGCATTTCTCAAGAGACAATTCGAAGGCATATAGATACAGCTGCCCTGCCACG 626  
Db 1078 TGCAGGCCCTTCTCAAGAGAGATTCGAAGGCATATAGTGTGTCCAGCCACC 1137  
QY 627 AATGAATGTGAATCACAAGCCGACAGCTAATCTCTGCCAGGCTTGCCGCTCATGAG 686  
Db 1138 AACCAAGTCAACCATGTATATAAACAAGAGGAAGAGCTGCCAGGCTGCCGCAAA 1197  
QY 687 TGTTTAAAAGTGGGCATGTGAAAAGAGGGGTGCGTCTTGACAGAGTACGTGAGGTCGG 746  
Db 1198 TGCTACGAAGTGGGAATGATGAAGGTGGGATACGAAAAAGACGAAGAGAGGAGAAATG 1257  
QY 747 CAGAAGTACAAGCGCAGATAGTGGGAGAA----- 778  
Db 1258 TTGAACACACAAGCCAGAGAGATGATGGGAGGCGAGGGGTGAAGTGGGCTGTGCTGA 1317  
QY 779 -----CAGCCCATACCTGAACCCCTCAGCTGCTGACCCAGCCCAAAAA----- 820  
Db 1318 GACATGAGAGCTGCCAACCTTTGGCCAAAGCCGCTCATGATCAACGCTTAAGAAGAAC 1377  
QY 821 -----GCCATATACAAGATTTGTCTCAACATTTGTTGGCTGACCG 863  
Db 1378 AGCCTGGCCTTGTCCCTGACGGCCGACAGATGTCATGGCCTTGTGGATGCTGAGCCC 1437  
QY 864 GAGAAGATCTATGCCATGCTGACCTGACCTACTGTCCCGACAGTGAATCAAGCCCTCACT 923  
Db 1438 CCCATACTCTATTCGAGTATGATCTACCAAGACCCCTTCAGTGAAGCTTCGATGATGGGC 1497  
QY 924 ACACTGTGTGACTTGGCCGACGAGTGTGTTATCATTTGATGGGCGAAGCATATT 983  
Db 1498 TTACTGACCAACCTGGCAGACAGGAGCTGGTTCAACATGATCAACTGGGCGAAGAGGTTG 1557  
QY 984 CCAGGCTTCTCCAGCTGTCCCTGGCGACCAATGAGCCTTTCGAGAGTGTCTTGATG 1043  
Db 1558 CCAGGCTTTTGGAATTTGACCCCTCCATGATCAGGTCCACCTTCAAGATGTGCTGCTA 1617  
QY 1044 GAAATTTGATCCTTGGTGTGATACCGGCTCTTTTCATTTGAGATGAACCTGTCTAT 1103  
Db 1618 GAGATCCTGATGATGTGCTCTGCTGGCGCTTCATGGAGCACCCAAGTAAAGCTACTGTTT 1677  
QY 1104 GCAGACGATATATATAGACGAAGACCAAGCTCAAAAT---TAGCAGGCTTCTTGATCTA 1160  
Db 1678 GCTCCTTAACCTTCTTGTGACAGGAACCAAGGAAAAATGTGTAGAGGGCATGTTGAGATC 1737

QY 1161 AATAATGCTATCTGCAGCTGTGTAAGAATAACAAGACATGAAGCTGAAAAAGAGAA 1220  
Db 1738 TTCGACATGCTGTGCTACATCATCTCGGTTCCGATGATGAATCTGCAGGAGAGAG 1797  
QY 1221 TTTGTCAACCTGAAGCTATAGCTCTTGCTAATTGAG 1257  
Db 1798 TTTGTGTGCTCAAAATCTATATTTTGCTTAATTCTG 1834

RESULT 9

US-08-481-970-9  
Sequence 9, Application US/08481970  
Patent No. 5859310  
GENERAL INFORMATION:  
APPLICANT: Gossen, Manfred  
APPLICANT: Bujard, Hermann  
APPLICANT: Salfeld, Jochen  
APPLICANT: Voss, Jeffrey  
TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lahive & Cockfield  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,970  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/260,452  
FILING DATE: 14-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/076,327

FILING DATE: 14-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Giulio A. DeConti, Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: BBI-013CP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4963 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Human cytomegalovirus  
IMMEDIATE SOURCE:  
CLONE: PUHD BGR4  
US-08-481-970-9

Query Match 4.6%; Score 136.6; DB 2; Length 4963;  
Best Local Similarity 52.3%; Pred. No. 7.8e-27;  
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

QY 507 TGTTTAGTGTGTGACATCGCTTCTGGGTACCATATAGGGGTAGCATCATGTGAAGC 566  
Db 1018 TGTGCAGTGTGCAATGACTATGCTTCAAGCTACCATATAGAGTCTGTCTGTGAGGC 1077  
QY 567 TGCAGGCATTTCTCAAGAGACAATTCGAAGGCATATAGATACAGCTGCCCTGCCACG 626  
Db 1078 TGCAGGCCCTTCTCAAGAGAGATTCGAAGACATACGACTATATGTGTCCAAGCCACC 1137







STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/041,886  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 2626  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6450 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 361..2146  
US-09-041-886-34

Query Match 4.6%; Score 136.6; DB 3; length 6450;  
Best Local Similarity 52.3%; Pred. No. 9.1e-27;  
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

QY 507 TGTTTAGTGTGTGTGACATCGCTTCTGGGTACCACTATGGGGTAGCATGTGAAGCC 566  
DB 913 TGTGAGTGTGCAATGACTATGCTTCAGGCTACCATTTAGAGTCTGTCTGTGAGGGC 972  
QY 567 TGCAGGCAATCTTCAAGAGCAATTCAAGGCAATATAGATACAGCTGCCCTGCCAGC 626  
DB 973 TGCAGGCTTCTTCAAGAGAGTATTCAAGGACATACATATATGTGTCCAGCCACC 1032  
QY 627 AATGAATGTGAATCACAAGCGCAGACGTAATCTGCCAGGCTTGGCGCTTCATGAAG 686  
DB 1033 AACCAAGTGCACCATTTGATATAAAACAGAGAGAGAGCTGCCAGGCTGCCGCAAA 1092  
QY 687 TGTTTAAAGTGGCATGCTGAAGAAGGGGTGCTCTTGAACAGATACGTGGAGGTCCG 746  
DB 1093 TGCTACGAAGTGGGAATGATGAAGGTGGGATACGAAAAAGACCGAAGAGAGGAGATG 1152  
QY 747 CAGAAGTACAAGCGCAGGATAGATCGCGAGAA----- 778  
DB 1153 TTGAAACACAAGCGCCAGAGAGATGATGGGAGGGGCAAGGGGTGAAGTGGGGTCTGTGGA 1212  
QY 779 -----CAGCCCATACCTGAACCTTCAGCTGCTTCAGCCAGCCAAAA----- 820  
DB 1213 GACATGAGAGCTGCCAACCTTTGGCCCAAGCCCGCTCATGATCAACGCTCTAAGAGAAC 1272  
QY 821 -----GCCATATAACAAGATTGTCTACATTTGTTGGTGGTGAACCG 863  
DB 1273 AGCCTGGCCTTGTCCCTGACGCGCCAGCAGATGGTCACTGCTTGTGATGCTGAGCCC 1332  
QY 864 GAGAAGATCTATGCGCATGCGCTGACCCCTACTGTCCCGACAGTGAATCAAGCCCTCACT 923  
DB 1333 CCCATATCTATTCGAGTATGATCCTTACAGACCCCTTCACTGAAGCTTCATGATGGGC 1392  
QY 924 ACACTGTGTGACTTGGCCGACCGAGAGTGGTGTATCATTTGATGGGGCAAGCATATT 983  
DB 1393 TTACTGACCAACCTGGCAGACAGGAGCTGGTTCATCACTGAGGCGGAGAGGGGTG 1452  
QY 984 CCAGGCTTCTCCACGCTGTCCCTGGCGGACAGATGAGCCTTCTGACAGTGTGATG 1043  
|||||

DB 1453 CCAGGCTTGTGATTTGACCCCTCCATGATCAGGTCCACCCTTCTAGAATGTGCTGCTA 1512  
QY 1044 GAAATTTGATCCCTTGTGTGCTGATATACCGGTCTTTCATTTGAGAGTGAATGTCTAT 1103  
DB 1513 GAGATCCTGATGATTTGCTGCTGCTGGCGCTCCATGAGACCCAGTGAAGCTACTGTT 1572  
QY 1104 GCAGACGATTAATAATGAGCAGGAGACCAAT--TAGAGGCTTCTGTATCTA 1160  
DB 1573 GCTCCTAACTTCTCTTGGACAGAAACAGGAAATGTGTAGAGGCGATGTGAGATC 1632  
QY 1161 AATAATGCTATCTCGACGCTGTAAAGAAATACAAAGACATGAAGCTGAAAAAGAGAA 1220  
DB 1633 TTGACATGCTGCTGCTGCTACATCATCTCGGTTCCGATGATGAATCTGCAAGAGAGAG 1692  
QY 1221 TTGTCAACCTCAAGCTATAGCTCTTGTCTAATTCAG 1257  
DB 1693 TTGTGTGCTCAATCTATTTTGTCTTAATTCTG 1729

RESULT 13

US-08-453-998-1  
Sequence 1, Application US/08453998  
Patent No. 644438

GENERAL INFORMATION:

APPLICANT: CHAMON, PIERRE  
APPLICANT: METZGER, DANIEL  
APPLICANT: WHITE, JOHN  
TITLE OF INVENTION: METHOD FOR THE PREPARATION OF A PROTEIN  
TITLE OF INVENTION: BY YEASTS USING AN INDUCIBLE SYSTEM, VECTORS AND  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1100 NEW YORK AVE., N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,998  
FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/161,064  
FILING DATE: 03-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: CHAPIN, MARIANA K.  
REGISTRATION NUMBER: 35,843  
REFERENCE/DOCKET NUMBER: 1037/98493  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-861-3711  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 6450 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-453-998-1

Query Match 4.6%; Score 136.6; DB 4; length 6450;  
Best Local Similarity 52.3%; Pred. No. 9.1e-27;  
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

QY 507 TGTTAGTGTGTGACATCGCTTCTGGGTACCACTATGGGGTAGCATCATGTGAAGCC 566  
|||||

```
Db      913 TGTGCAGTGTGCAATGACTATGCTTCAGGCTACCATATGAGTCTGCTCCTGTGAGGC 972
QY      567 TGCAGGCAATTCTTCAAGAGACAATTCAAGGCAATATAGAAATACAGCTGCCCTGCCACG 626
Db      973 TGCAGGCAATTCTTCAAGAGACAATTCAAGGCAATATAGAAATACAGCTGCCCTGCCACG 1032
QY      627 AATGAATGTGAATCAACAAGCGCAGACGTAATCCTGCCAGGCTTGCCGCTTCATGAAG 686
Db      1033 AACCAAGTGCACCAATTGATATAAAACAGAGAGAAAGAGCTGCCAGGCTTGCCGCTGCCAAA 1092
QY      687 TGTTTAAAGTGGGCAATGCTGAAGAAGGGGTGCGTCTTGACAGAGTACGTGAGGTCCG 746
Db      1093 TGCTACGAGTGGGAATGATGAAGAGTGGGATACGAAAAAGACCGAAGAGAGAGAGAAATG 1152
QY      747 CAGAGTACAAAGCCGAGATAGATGCGGAGAA----- 778
Db      1153 TTGAAACACAAAGCCGACAGAGATGATGGGAGGGGCAAGGGGTGAAGTGGGCTCTGCTGA 1212
QY      779 -----CAGCCCATACCTGAACCTTCAGCTGGTTCAGCCCAAAAA----- 820
Db      1213 GACATGAGAGCTGCCAACCTTTGGCCAAAGCCGCTCATGATCAAAACGCTTAAGAAAGAAC 1272
QY      821 -----GCCATATAACAAGATTGTCTACATTTGTTGGTGGCTGAACCG 863
Db      1273 AGCCTGGCCTTGTCTCCCTGACGGCCGACAGATGTCAGTCCCTTGTGATGCTGAACCC 1332
QY      864 GAGAAGATCTATGCGCATGCTGACCTGCTACTGTCCCGACAGTGAATCAAAAGCCCTCACT 923
Db      1333 CCCATACTCTATTCGAGTATGATCCTACAGAACCTTCACTGAAGCTTCGATGATGGGC 1392
QY      924 ACACGTGTGACTTGGCCGACCGAGAGTGTGTGTTATCATTTGATGGGCGAAGCATATT 983
Db      1393 TTAAGTACCAACCTGGCAGACAGGAGAGCTGTTCAATGATCAACTGGCGAAGAGGCTG 1452
QY      984 CCAGGCTTCTCCACGCTGTCCCTTGCGCGGACCAAGATGAGCCTTCTGACAGTGGCTGATG 1043
Db      1453 CCAAGCTTTGTGATTTGACCTTCCATGATCAAGGTCCACCTTCTAGAATGTGCTGGCTA 1512
QY      1044 GAAATTTGATCCTTGGTGTGTATACCGGTCTCTTCAATTTGAGATGAATGTGTCTAT 1103
Db      1513 GAGATCCTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1572
QY      1104 GCAGACGATTATATATATGACGAAGAACCAAGTCCAAAT---TAGCAGGCTTCTTGAATCTA 1160
Db      1573 GCTCCTAACTTGTCTCTTGACAGAAACCAAGGAAATGTGTAGAGGCGATGTGAGATC 1632
QY      1161 AATAATGCTATCTGACAGCTGTGAAGAAATACAAAGACATGAAGCTGAAAAAGAGAA 1220
Db      1633 TTGCACATGCTGCTGCTACATCATCTCGGTTCCGATGATGAATCTGACAGGAGAGAG 1692
QY      1221 TTTGTACCCCTCAAAGCTATAGCTCTTGTATTCTAG 1257
Db      1693 TTTGTGTGCTCAAAATCTATTTTGTCTTAATTCTG 1729

RESULT 14
US-10-027-983-3
; Sequence 3, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RFS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 3
; LENGTH: 6450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (361)...(2148)
US-10-027-983-3
Query Match      4.6%; Score 136.6; DB 4; Length 6450;
Best Local Similarity 52.3%; Pred. No. 9.1e-27;
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

QY      507 TGTTTAGTGTGTGTGACATGCTTCTGGGTAACCACTATGGGCTAGCATCATGTGAAGCC 566
Db      913 TGTGAGTGTGCAATGACTATGCTTCAAGGCTAACATATGAGTGTGTGTGTGTGTGTGT 972
QY      567 TGCAGGCAATTCTTCAAGAGACAATTCAAGGCAATATAGAAATACAGCTGCCCTGCCACG 626
Db      973 TGCAGGCAATTCTTCAAGAGACAATTCAAGGCAATATAGAAATACAGCTGCCCTGCCACG 1032
QY      627 AATGAATGTGAATCAACAAGCGCAGACGTAATCCTGCCAGGCTTGCCGCTTCATGAAG 686
Db      1033 AACCAAGTGCACCAATTGATATAAAACAGAGAGAAAGAGCTGCCAGGCTTGCCGCTGCCAAA 1092
QY      687 TGTTTAAAGTGGGCAATGCTGAAGAAGGGGTGCGTCTTGACAGAGTACGTGAGGTCCG 746
Db      1093 TGCTACGAGTGGGAATGATGAAGAGTGGGATACGAAAAAGACCGAAGAGAGAGAGAAATG 1152
QY      747 CAGAGTACAAAGCCGAGATAGATGCGGAGAA----- 778
Db      1153 TTGAAACACAAAGCCGACAGAGATGATGGGAGGGGCAAGGGGTGAAGTGGGCTCTGCTGA 1212
QY      779 -----CAGCCCATACCTGAACCTTCAGCTGGTTCAGCCCAAAAA----- 820
Db      1213 GACATGAGAGCTGCCAACCTTTGGCCAAAGCCGCTCATGATCAAAACGCTTAAGAAAGAAC 1272
QY      821 -----GCCATATAACAAGATTGTCTACATTTGTTGGTGGCTGAACCG 863
Db      1273 AGCCTGGCCTTGTCTCCCTGACGGCCGACAGATGTCAGTCCCTTGTGATGCTGAACCC 1332
QY      864 GAGAAGATCTATGCGCATGCTGACCTGCTACTGTCCCGACAGTGAATCAAAAGCCCTCACT 923
Db      1333 CCCATACTCTATTCGAGTATGATCCTACAGAACCTTCACTGAAGCTTCGATGATGGGC 1392
QY      924 ACACGTGTGACTTGGCCGACCGAGAGTGTGTGTTATCATTTGATGGGCGAAGCATATT 983
Db      1393 TTAAGTACCAACCTGGCAGACAGGAGAGCTGTTCAATGATCAACTGGGCGAAGAGGCTG 1452
QY      984 CCAGGCTTCTCCACGCTGTCCCTTGCGCGGACCAAGATGAGCCTTCTGACAGTGGCTGATG 1043
Db      1453 CCAAGCTTTGTGATTTGACCTTCCATGATCAAGGTCCACCTTCTAGAATGTGCTGGCTA 1512
QY      1044 GAAATTTGATCCTTGGTGTGTATACCGGTCTCTTCAATTTGAGATGAATGTGTCTAT 1103
Db      1513 GAGATCCTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1572
QY      1104 GCAGACGATTATATATATGACGAAGAACCAAGTCCAAAT---TAGCAGGCTTCTTGAATCTA 1160
Db      1573 GCTCCTAACTTGTCTCTTGACAGAAACCAAGGAAATGTGTAGAGGCGATGTGAGATC 1632
QY      1161 AATAATGCTATCTGACAGCTGTGAAGAAATACAAAGACATGAAGCTGAAAAAGAGAA 1220
Db      1633 TTGCACATGCTGCTGCTACATCATCTCGGTTCCGATGATGAATCTGACAGGAGAGAG 1692
QY      1221 TTTGTACCCCTCAAAGCTATAGCTCTTGTATTCTAG 1257
Db      1693 TTTGTGTGCTCAAAATCTATTTTGTCTTAATTCTG 1729

RESULT 15
US-09-833-381-269/C
; Sequence 269, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2004, 23:22:43 ; Search time 2392 Seconds  
(without alignments)  
6118.676 Million cell updates/sec

Title: US-10-054-841-3  
Perfect score: 2985  
Sequence: 1 gcggcgccgagtggtgtg.....actggcgccgcgcttgat 2985

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues  
Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2985	100.0	2985	15 US-10-054-841-3	Sequence 3, Appli
2	2973	99.6	2987	15 US-10-054-841-5	Sequence 5, Appli
3	1377	46.1	1377	16 US-10-355-218-3	Sequence 3, Appli
4	715	24.0	2807	15 US-10-054-841-1	Sequence 1, Appli
5	695	23.3	2153	15 US-10-101-510-87	Sequence 87, Appli
6	695	23.3	2153	15 US-10-101-510-413	Sequence 413, Appli
7	428	14.3	622	15 US-10-054-841-8	Sequence 8, Appli
8	424.4	14.2	2218	9 US-09-962-436-268	Sequence 268, Appli
9	424.4	14.2	2218	17 US-10-776-827-4	Sequence 4, Appli
10	421.4	14.1	2539	15 US-10-037-270-454	Sequence 454, Appli
11	421.4	14.1	2539	16 US-10-117-722-454	Sequence 454, Appli
12	421.4	14.1	4250	16 US-10-297-653-2	Sequence 2, Appli
13	401.6	13.5	2402	13 US-10-342-887-1041	Sequence 1041, Appli
14	401.6	13.5	2402	13 US-10-172-118-1041	Sequence 1041, Appli

15	401.6	13.5	2402	15 US-10-211-239-7	Sequence 7, Appli
16	401.6	13.5	2402	15 US-10-101-510-66	Sequence 66, Appli
17	360.4	12.1	1599	16 US-10-093-463-201	Sequence 201, Appli
18	204.6	6.9	403	15 US-10-054-841-7	Sequence 7, Appli
19	191	6.4	565	13 US-10-027-632-285776	Sequence 285776, Appli
20	191	6.4	565	16 US-10-027-632-285776	Sequence 285776, Appli
21	189.8	6.4	565	13 US-10-027-632-285777	Sequence 285777, Appli
22	189.8	6.4	565	16 US-10-027-632-285777	Sequence 285777, Appli
23	141.4	4.7	1978	15 US-10-052-092-23	Sequence 23, Appli
24	141.4	4.7	1978	15 US-10-437-107-23	Sequence 23, Appli
25	136.6	4.6	1785	15 US-10-095-373A-1	Sequence 1, Appli
26	136.6	4.6	1788	9 US-09-853-033-1	Sequence 1, Appli
27	136.6	4.6	2092	10 US-09-952-680A-9	Sequence 9, Appli
28	136.6	4.6	2092	15 US-10-052-092-6	Sequence 6, Appli
29	136.6	4.6	2092	15 US-10-157-899A-54	Sequence 54, Appli
30	136.6	4.6	2092	15 US-10-437-107-6	Sequence 6, Appli
31	136.6	4.6	2106	15 US-10-052-092-5	Sequence 5, Appli
32	136.6	4.6	2106	15 US-10-437-107-5	Sequence 5, Appli
33	136.6	4.6	4963	9 US-09-281-674-9	Sequence 9, Appli
34	136.6	4.6	4963	9 US-09-892-227-9	Sequence 9, Appli
35	136.6	4.6	4963	13 US-09-777-317-9	Sequence 9, Appli
36	136.6	4.6	6450	13 US-10-342-887-435	Sequence 435, Appli
37	136.6	4.6	6450	13 US-10-172-118-435	Sequence 435, Appli
38	136.6	4.6	6450	13 US-10-388-360-303	Sequence 303, Appli
39	136.6	4.6	6450	13 US-10-448-753-3	Sequence 3, Appli
40	136.6	4.6	6450	14 US-10-096-710-2	Sequence 2, Appli
41	136.6	4.6	6450	14 US-10-081-563-1	Sequence 1, Appli
42	136.6	4.6	6450	15 US-10-052-092-1	Sequence 1, Appli
43	136.6	4.6	6450	15 US-10-052-092-7	Sequence 7, Appli
44	136.6	4.6	6450	15 US-10-207-655-60	Sequence 60, Appli
45	136.6	4.6	6450	15 US-10-177-293-127	Sequence 127, Appli

ALIGNMENTS

RESULT 1  
US-10-054-841-3  
; Sequence 3, Application US/10054841  
; Publication No. US20030119100A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Fang  
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
; FILE REFERENCE: 19999YCA  
; CURRENT APPLICATION NUMBER: US/10/054,841  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: 09/487,379  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: 09/141,000  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/078,633  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/062,902  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/057,090  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2985  
; TYPE: DNA  
; ORGANISM: Human  
; US-10-054-841-3

Query Match 100.0%; Score 2985; DB 15; Length 2985;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GC GGCGCCGCGAGTGTGTGAATTCGGCTGTCACTAGAGAACATTTGTTAATTGCA 60  
DB 1 GC GGCGCCGCGAGTGTGTGAATTCGGCTGTCACTAGAGAACATTTGTTAATTGCA 60

QY 61 CTGTGCTCTGTCAAGGAAACTTTGATTATAGCTGGGGTGACAAATAATGTTGCCGGT 120  
Db 61 CTGTGCTCTGTCAAGGAAACTTTGATTATAGCTGGGGTGACAAATAATGTTGCCGGT 120  
QY 121 CGCACATGGATTGGGTAGAACTTTGCCCTTGAACTTTTCCCTGCACTACGAGGAAG 180  
Db 121 CGCACATGGATTGGGTAGAACTTTGCCCTTGAACTTTTCCCTGCACTACGAGGAAG 180  
QY 181 AGCTTCTGTGAGAAATGTCAAAACAAGATGCACATTGATTCAGCTGTTCCTTCA 240  
Db 181 AGCTTCTGTGAGAAATGTCAAAACAAGATGCACATTGATTCAGCTGTTCCTTCA 240  
QY 241 TCAAGACGGAACTTCCAGCCCCAGCCCTCCCTGACGGACAGCGTCAACCAACAGCCCTG 300  
Db 241 TCAAGACGGAACTTCCAGCCCCAGCCCTCCCTGACGGACAGCGTCAACCAACAGCCCTG 300  
QY 301 GTGGCTCTTACAGACGCCAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGAC 360  
Db 301 GTGGCTCTTACAGACGCCAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGAC 360  
QY 361 TTGACTTCGCCACTCTCTACCCCTTCTGCTCCTATCCTGGGAGGTAGTGGCCCTGTACGA 420  
Db 361 TTGACTTCGCCACTCTCTACCCCTTCTGCTCCTATCCTGGGAGGTAGTGGCCCTGTACGA 420  
QY 421 AACTGTATGATGACTGCTCTCCAGCACCATGTGTGAAGATCCCCAGACCAAGTGTGAATACA 480  
Db 421 AACTGTATGATGACTGCTCTCCAGCACCATGTGTGAAGATCCCCAGACCAAGTGTGAATACA 480  
QY 481 TGTCTAATCTGATGCCCCAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACC 540  
Db 481 TGTCTAATCTGATGCCCCAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACC 540  
QY 541 ACTATGGGGTAGCATCATGTGAAGCCTGCAGAGCATTTCTTCAAGAGACAAATTCAGGGCA 600  
Db 541 ACTATGGGGTAGCATCATGTGAAGCCTGCAGAGCATTTCTTCAAGAGACAAATTCAGGGCA 600  
QY 601 ATATAGAAATACAGCTGCCCTGCCCGCAAGATGAATGTGAATCACAAGCGCAGACGTAAAT 660  
Db 601 ATATAGAAATACAGCTGCCCTGCCCGCAAGATGAATGTGAATCACAAGCGCAGACGTAAAT 660  
QY 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAAAGAGGGGTGC 720  
Db 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAAAGAGGGGTGC 720  
QY 721 GTCTTGACAGAGTACGTGAGGTGGCAGAAGTACAAGCGCAGATAGATGCGAGACA 780  
Db 721 GTCTTGACAGAGTACGTGAGGTGGCAGAAGTACAAGCGCAGATAGATGCGAGACA 780  
QY 781 GCCCATACCTGAACCCCTCAGCTGTTCAGCCAGCCAAAAGCCATATAACAAGATTGTCT 840  
Db 781 GCCCATACCTGAACCCCTCAGCTGTTCAGCCAGCCAAAAGCCATATAACAAGATTGTCT 840  
QY 841 CACATTTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGACCCCTACTGTCCCG 900  
Db 841 CACATTTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGACCCCTACTGTCCCG 900  
QY 901 ACAGTGACATCAAAAGCCCTCACTACACTGTGTGACTTTGGCCGACCGAGATTGGTGA 960  
Db 901 ACAGTGACATCAAAAGCCCTCACTACACTGTGTGACTTTGGCCGACCGAGATTGGTGA 960  
QY 961 TCATTGGATGGGCGAAGCATATTCAGGCTTCTCCAGCGTGTCCCTGGCGACCAAGTGA 1020  
Db 961 TCATTGGATGGGCGAAGCATATTCAGGCTTCTCCAGCGTGTCCCTGGCGACCAAGTGA 1020  
QY 1021 GCCTTCTGCAGAGTGTGGATGGAATTTTGAATCCCTGGTGTGCTATACCGGTCTCTT 1080  
Db 1021 GCCTTCTGCAGAGTGTGGATGGAATTTTGAATCCCTGGTGTGCTATACCGGTCTCTT 1080  
QY 1081 CATTTGAGGATGAACCTTGTCTATGACAGACGATTATATGAGACGAAGACCAAGTCCAAT 1140  
Db 1081 CATTTGAGGATGAACCTTGTCTATGACAGACGATTATATGAGACGAAGACCAAGTCCAAT 1140  
QY 1141 TAGCAGGCTTCTTGTACTAAATAATGCTATCCTGCAGCTGTTAAAGAAATACAGAGCA 1200

Db 1141 TAGCAGGCTTCTTGTATCTAAATAATGCTATCCTGCAGCTGTTAAAGAAATACAGAGCA 1200  
QY 1201 TGAAGCTGGA AAAAAGAAATTTGTACACCTCAAAAGCTATAGCTCTGTAAATTCAGACT 1260  
Db 1201 TGAAGCTGGA AAAAAGAAATTTGTACACCTCAAAAGCTATAGCTCTGTAAATTCAGACT 1260  
QY 1261 CCATGCACATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAGGATGTCTTACATGAAGCGC 1320  
Db 1261 CCATGCACATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAGGATGTCTTACATGAAGCGC 1320  
QY 1321 TGCAAGATTATGAAGCTGGCCAGCATGGAAGACCCCTCGTCGAGCTGGCAAGATGCTGA 1380  
Db 1321 TGCAAGATTATGAAGCTGGCCAGCATGGAAGACCCCTCGTCGAGCTGGCAAGATGCTGA 1380  
QY 1381 TGACACTGCCACTCCTGAGGCGACCTCTACCAAGGCCGTGACGATTTCTACACATCA 1440  
Db 1381 TGACACTGCCACTCCTGAGGCGACCTCTACCAAGGCCGTGACGATTTCTACACATCA 1440  
QY 1441 AACTAGAGGCAAAAGTCCCAATGCACAAACCTTTTTTGGAAATGTGAGGCCAAGGTCT 1500  
Db 1441 AACTAGAGGCAAAAGTCCCAATGCACAAACCTTTTTTGGAAATGTGAGGCCAAGGTCT 1500  
QY 1501 GACTAAAAGCTCCCTGGGCTTCCCATCTTCATGTTGAAAAGGGAATAAACCCAAAG 1560  
Db 1501 GACTAAAAGCTCCCTGGGCTTCCCATCTTCATGTTGAAAAGGGAATAAACCCAAAG 1560  
QY 1561 AGTGATGTCGAAGAACTTAGATTAGTTAACAACATCAAAAATCAACAGACTGCATG 1620  
Db 1561 AGTGATGTCGAAGAACTTAGATTAGTTAACAACATCAAAAATCAACAGACTGCATG 1620  
QY 1621 ATAAATTTAGCAGCAAGACTATGAAGCAGCTTTCAGATTCTCTCAATAGGTTCTGATGACT 1680  
Db 1621 ATAAATTTAGCAGCAAGACTATGAAGCAGCTTTCAGATTCTCTCAATAGGTTCTGATGACT 1680  
QY 1681 TCTTTCTACTTTCTCCATCATCTTTCTTCTCTCTTCCACATTTCTCTTCTCTTTA 1740  
Db 1681 TCTTTCTACTTTCTCCATCATCTTTCTTCTCTCTTCCACATTTCTCTTCTCTTTA 1740  
QY 1741 TTTTCTCTCTTTCTTCTTCTTCACTCCCTTATTTCTTGTCTTTCAATTCCTAGTTCC 1800  
Db 1741 TTTTCTCTCTTTCTTCTTCTTCACTCCCTTATTTCTTGTCTTTCAATTCCTAGTTCC 1800  
QY 1801 CATCTCCTTTATTTTCTTCCCGTCTGCGCTGCGCTTCTTCTTTCTTACCTACTCTCAT 1860  
Db 1801 CATCTCCTTTATTTTCTTCCCGTCTGCGCTGCGCTTCTTCTTTCTTACCTACTCTCAT 1860  
QY 1861 TCCCTCTTTTCTCATCTCTTCCCTTTTCTTAATTTGAATAAGCTTTAGTTAAABAA 1920  
Db 1861 TCCCTCTTTTCTCATCTCTTCCCTTTTCTTAATTTGAATAAGCTTTAGTTAAABAA 1920  
QY 1921 AAAAATCCTCCCTTCCCGCTTTCCTTTCCCTTCTTCCCTTTTCCCTTTTCCCT 1980  
Db 1921 AAAAATCCTCCCTTCCCGCTTTCCTTTTCCCTTCTTCCCTTTTCCCTTTTCCCT 1980  
QY 1981 TTCCTTTCTTCTTCTTGAACCTTCTTCCATCTTTCTTTTCTTCTTCTGCTGTGA 2040  
Db 1981 TTCCTTTCTTCTTCTTGAACCTTCTTCCATCTTTCTTTTCTTCTTCTGCTGTGA 2040  
QY 2041 CTTTAAAGAGGTCTTAACTGAAGAGATGGAAGCCAGCCCTGCCAAAGAGATGAGA 2100  
Db 2041 CTTTAAAGAGGTCTTAACTGAAGAGATGGAAGCCAGCCCTGCCAAAGAGATGAGA 2100  
QY 2101 TCCATATATGATGCGCAGTGAACCTTATTTGAACCATACCGTCCCAATGACTAAGAA 2160  
Db 2101 TCCATATATGATGCGCAGTGAACCTTATTTGAACCATACCGTCCCAATGACTAAGAA 2160  
QY 2161 TCAAGAGAGAGAAACCAAGTTCCTAAAGTACAGTGCACATATACAAATTGACTGACT 2220  
Db 2161 TCAAGAGAGAGAAACCAAGTTCCTAAAGTACAGTGCACATATACAAATTGACTGACT 2220  
QY 2221 GCAGTATTAGATTTCATGGAGAGCAGCTCTAATTAGCAACTTAAGCAACGTTGACTGG 2280

Db 2221 GCAGTATTAGATTTCATGGAGACAGCCTCTAAATTAGACACTTAAGCAAGTTGCATCGG 2280  
QY 2281 CTGCTTCTTATTCATTGCTTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTAATTGT 2340  
Db 2281 CTGCTTCTTATTCATTGCTTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTAATTGT 2340  
QY 2341 TTTTCAAGTCTTCCAGGTAATTTGTTAGTTAGCTACTATGTAACTTTTTCAGGGGAATAG 2400  
Db 2341 TTTTCAAGTCTTCCAGGTAATTTGTTAGTTAGCTACTATGTAACTTTTTCAGGGGAATAG 2400  
QY 2401 TTTAAGCTTTATTCATTCAATGCAATACTAAAGAGAAATAAGAAATGCAATTTTGTGCT 2460  
Db 2401 TTTAAGCTTTATTCATTCAATGCAATACTAAAGAGAAATAAGAAATGCAATTTTGTGCT 2460  
QY 2461 GGCTTTGAACAATTAACGAACAATAATGAAGAGACAATGATCCTGAAGGAGATTTTAA 2520  
Db 2461 GGCTTTGAACAATTAACGAACAATAATGAAGAGACAATGATCCTGAAGGAGATTTTAA 2520  
QY 2521 AAATGTTTGTCTTCTTCAAAATGGAGATTTTGTGTAACAGCTTTACCACTTTTCAG 2580  
Db 2521 AAATGTTTGTCTTCTTCAAAATGGAGATTTTGTGTAACAGCTTTACCACTTTTCAG 2580  
QY 2581 CCATTTATTAATATGGGAATTTAACTTACTCAAGCAATAGTTGAAGGGAAGTGCATATT 2640  
Db 2581 CCATTTATTAATATGGGAATTTAACTTACTCAAGCAATAGTTGAAGGGAAGTGCATATT 2640  
QY 2641 ATCAAGGATGCAATTTATGTTGTGTCAGAGTCTGTGCCAAACATCAATTTCTTAACATG 2700  
Db 2641 ATCAAGGATGCAATTTATGTTGTGTCAGAGTCTGTGCCAAACATCAATTTCTTAACATG 2700  
QY 2701 AGCTCCAGTTTACCTAAATGTTCACTGACACAAGGATGAGATTACACCTACAGTGACTC 2760  
Db 2701 AGCTCCAGTTTACCTAAATGTTCACTGACACAAGGATGAGATTACACCTACAGTGACTC 2760  
QY 2761 TGAGTAGTCACATATATATAAGCACTGCACATGAGATATAGTCCGTAGAATTGTCAAGAGT 2820  
Db 2761 TGAGTAGTCACATATATATAAGCACTGCACATGAGATATAGTCCGTAGAATTGTCAAGAGT 2820  
QY 2821 GCACCTCTCTACTTGGGAGGTACAAATGCCATATGATTTCTAGCTGCCATGTGTTAGG 2880  
Db 2821 GCACCTCTCTACTTGGGAGGTACAAATGCCATATGATTTCTAGCTGCCATGTGTTAGG 2880  
QY 2881 AATGTGATACTGCTGTTTGCAAAAGTCAAGACCTTGCTCGAAGGAGCTGTGAGCCAG 2940  
Db 2881 AATGTGATACTGCTGTTTGCAAAAGTCAAGACCTTGCTCGAAGGAGCTGTGAGCCAG 2940  
QY 2941 TATTCATTTAAGAGAATTCACCACACTGGCGCGCGCTTGAT 2985  
Db 2941 TATTCATTTAAGAGAATTCACCACACTGGCGCGCGCTTGAT 2985

RESULT 2  
US-10-054-841-5  
; Sequence 5, Application US/10054841  
; Publication No. US20030119100A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Fang  
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
; TITLE OF INVENTION: RECEPTOR PROTEINS  
; FILE REFERENCE: 19999YCA  
; CURRENT APPLICATION NUMBER: US/10/054, 841  
; PRIOR FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: 09/487, 379  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: 09/141, 000  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/078, 633  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/062, 902  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/057, 090  
; PRIOR FILING DATE: 1997-08-27  
; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2987  
; TYPE: DNA  
; ORGANISM: Human  
US-10-054-841-5  
Query Match 99.6%; Score 2973; DB 15; Length 2987;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2985; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
QY 1 GCGGGCCGCCAGTGTGTGAATTCGGCTTGTCACTAGAGAAACATTTGTGTTAATTGCA 60  
Db 1 GCGGGCCGCCAGTGTGTGAATTCGGCTTGTCACTAGAGAAACATTTGTGTTAATTGCA 60  
QY 61 CTGTGCTCTGTCAAGAAACTTTGATTTATAGCTGGGGTGCACAAATATGTTGCCGGT 120  
Db 61 CTGTGCTCTGTCAAGAAACTTTGATTTATAGCTGGGGTGCACAAATATGTTGCCGGT 120  
QY 121 CGCACATGATTCGGTAGAACTTTGCCCTTCTGAACTTTTCCCTGCACTACGAGGAG 180  
Db 121 CGCACATGATTCGGTAGAACTTTGCCCTTCTGAACTTTTCCCTGCACTACGAGGAG 180  
QY 181 AGCTTCTGCAAGATGTCAAAACAAGATCGACATGATTCAGCTGTTCCTTTCA 240  
Db 181 AGCTTCTGCAAGATGTCAAAACAAGATCGACATGATTCAGCTGTTCCTTTCA 240  
QY 241 TCAAGACGGAACCTTCCAGCCAGCCTCCTGACGGGACAGCTCAACCAACACAGCCCTG 300  
Db 241 TCAAGACGGAACCTTCCAGCCAGCCTCCTGACGGGACAGCTCAACCAACACAGCCCTG 300  
QY 301 GTGGCTCTTCAAGCCGAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGAC 360  
Db 301 GTGGCTCTTCAAGCCGAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGGAC 360  
QY 361 TTGACTCGCAACCTCTTACCTTCTGCTCTTATCTTGGGAGGTAGTGGCCTGTCAAGA 420  
Db 361 TTGACTCGCAACCTCTTACCTTCTGCTCTTATCTTGGGAGGTAGTGGCCTGTCAAGA 420  
QY 421 AACTGTATGATGACTGCTCCAGCAACCATTTGTAAGATCCCCAGCAAGTGTGAATACA 480  
Db 421 AACTGTATGATGACTGCTCCAGCAACCATTTGTAAGATCCCCAGCAAGTGTGAATACA 480  
QY 481 TGCTCACTCGATGCCCAAGAGACTGTGTTAGTGTGTGATCATGCTTCTGGGTACC 540  
Db 481 TGCTCACTCGATGCCCAAGAGACTGTGTTAGTGTGTGATCATGCTTCTGGGTACC 540  
QY 541 ACTATGGGGTAGCATCATGTGAAGCCTGCAAGGCACTTTCAAGAGCAATTCAGGCA 600  
Db 541 ACTATGGGGTAGCATCATGTGAAGCCTGCAAGGCACTTTCAAGAGCAATTCAGGCA 600  
QY 601 ATATAGAAATACAGCTGCCCTGCCACGAATGAATGTGAATCAAAAGCGCAGACGTAAT 660  
Db 601 ATATAGAAATACAGCTGCCCTGCCACGAATGAATGTGAATCAAAAGCGCAGACGTAAT 660  
QY 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAAGAGGGGTGC 720  
Db 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAAGAGGGGTGC 720  
QY 721 GTCTTGACAGAGTACGTGAGGTGCGCAGAAAGTACAAAGCGCAGATGATGCGGAGACA 780  
Db 721 GTCTTGACAGAGTACGTGAGGTGCGCAGAAAGTACAAAGCGCAGATGATGCGGAGACA 780  
QY 781 GCCCATACCTGAACCCCTCAGCTGTTCAAGCCAGCAAAAAGCCATATAACAAGATGTCT 840  
Db 781 GCCCATACCTGAACCCCTCAGCTGTTCAAGCCAGCAAAAAGCCATATAACAAGATGTCT 840  
QY 841 CACATTTGTTGGTGTGCTGAACCGGAGAAATCTATGCCATGCTGACCTACTGTCCCG 900  
Db 841 CACATTTGTTGGTGTGCTGAACCGGAGAAATCTATGCCATGCTGACCTACTGTCCCG 900  
QY 901 ACAGTGACATCAAAAGCCCTCACTACTGTGTGACTTGGCCGACGAGAGTTGGTGTTA 960  
Db 901 ACAGTGACATCAAAAGCCCTCACTACTGTGTGACTTGGCCGACGAGAGTTGGTGTTA 960



Db 901 ACAGTGACATCAAAAGCCCTCCTACTACTGTGTGACTTGGCCGACCGAGAGTTGGTGTTA 960  
QY 961 TCATTGGATGGGCGAAGCATATTCCAGGCTTCTCCACGCTGTCCCTGGCGGACCAGATGA 1020  
Db 961 TCATTGGATGGGCGAAGCATATTCCAGGCTTCTCCACGCTGTCCCTGGCGGACCAGATGA 1020  
QY 1021 GCCCTTCTGAGAGTGTCTTGATGGAATTGATCCCTGTGTGTGATATACCGGTCTCTTT 1080  
Db 1021 GCCCTTCTGAGAGTGTCTTGATGGAATTGATCCCTGTGTGTGATATACCGGTCTCTTT 1080  
QY 1081 CATTTGAGATGAACCTTGTCTATGACAGCATTAATATATATGAGAGAACCAAGTCCAAAT 1140  
Db 1081 CATTTGAGATGAACCTTGTCTATGACAGCATTAATATATATGAGAGAACCAAGTCCAAAT 1140  
QY 1141 TAGCAGGCTTCTTGATCTAATAATGCTATCCCTGCAGCTGTGTAAGAAATACAGAGCA 1200  
Db 1141 TAGCAGGCTTCTTGATCTAATAATGCTATCCCTGCAGCTGTGTAAGAAATACAGAGCA 1200  
QY 1201 TGAAGCTGGAATAAGAGATTTGTCAACCTCAAGCTATAGCTGTGCTAATTCAAGACT 1260  
Db 1201 TGAAGCTGGAATAAGAGATTTGTCAACCTCAAGCTATAGCTGTGCTAATTCAAGACT 1260  
QY 1261 CCATGCACTAGAAAGATGTTGAAGCCGTTCAGAGCTTCAGATGTCTTACATGAAGCGC 1320  
Db 1261 CCATGCACTAGAAAGATGTTGAAGCCGTTCAGAGCTTCAGATGTCTTACATGAAGCGC 1320  
QY 1321 TGCAGATTTATGAAGCTGGCCAGACATG--GAAGACCTCGTCGAGCTGGCAAGATGCT 1378  
Db 1321 TGCAGATTTATGAAGCTGGCCAGACATG--GAAGACCTCGTCGAGCTGGCAAGATGCT 1378  
QY 1379 GATGACACTGCCACTCCTGAGGCGACCTCTTACCAAGCCGTGAGCATTTCTACAACAT 1438  
Db 1381 GATGACACTGCCACTCCTGAGGCGACCTCTTACCAAGCCGTGAGCATTTCTACAACAT 1440  
QY 1439 CAAACTAGAAAGCAAAAGTCCCAATGACAACAACTTTTTTGGAAATGTTGAGGCCAAGGT 1498  
Db 1441 CAAACTAGAAAGCAAAAGTCCCAATGACAACAACTTTTTTGGAAATGTTGAGGCCAAGGT 1500  
QY 1499 CTGACTAAAGCTCCCTGGGCTTCCCATCTCTCATGTGTAAGAAAGGAAATTAACCCA 1558  
Db 1501 CTGACTAAAGCTCCCTGGGCTTCCCATCTCTCTTCTCCACATTTCTCTTCTCTT 1560  
QY 1559 AGAGTGATGTGGAAGAACTTAGAGTTAGTTAACAACATCAAAAAATCAACAGACTGCAC 1618  
Db 1561 AGAGTGATGTGGAAGAACTTAGAGTTAGTTAACAACATCAAAAAATCAACAGACTGCAC 1620  
QY 1619 TGATTAATTTAGCAGCAAGACTATGAAGCAGCTTTCAGATTCCTCCATAGGTTCTGTATGA 1678  
Db 1621 TGATTAATTTAGCAGCAAGACTATGAAGCAGCTTTCAGATTCCTCCATAGGTTCTGTATGA 1680  
QY 1679 GTTCTTTCTACTTCTCCATCATCTTCTTCTCTTCTCTTCTCCACATTTCTCTTCTCT 1738  
Db 1681 GTTCTTTCTACTTCTCCATCATCTTCTTCTCTTCTCTTCTCCACATTTCTCTTCTCTT 1740  
QY 1739 TATTTTTTCTCCTTTCTCTTCTTCCACCTCCCTATTTCTTGTCTTCTTCTTCTTCTAGTT 1798  
Db 1741 TATTTTTTCTCCTTTCTCTTCTTCCACCTCCCTATTTCTTGTCTTCTTCTTCTTCTAGTT 1800  
QY 1799 CCCATTTCTCCTTATTTTCTTCCCGTCTGCGCTGCGCTTCTTCTTCTTCTTCTTCTCTC 1858  
Db 1801 CCCATTTCTCCTTATTTTCTTCCCGTCTGCGCTGCGCTTCTTCTTCTTCTTCTTCTCTC 1860  
QY 1859 ATTCCTCTCTTTCTCATCTTCCCTTTTCTTAAATTTGAATAGCTTTAGTTAAAA 1918  
Db 1861 ATTCCTCTCTTTCTCATCTTCCCTTTTCTTAAATTTGAATAGCTTTAGTTAAAA 1920  
QY 1919 AAAAAAATCCTCCCTTCCCTTCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTC 1978  
Db 1921 AAAAAAATCCTCCCTTCCCTTCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTC 1980  
QY 1979 CTTTCTTCTTCTTCTTCTTGAACCTTCTTCCATCTTTCTTTTCTTCTTCTTCTTCTGCTG 2038  
Db 1981 CTTTCTTCTTCTTCTTCTTGAACCTTCTTCCATCTTTCTTTTCTTCTTCTTCTGCTGCTG 2040

QY 2039 AACTTTTAAAGAGGTCTCTAATCTGAAGAGATGGAAGCCAGCCCTGCCAAAGATGA 2098  
Db 2041 AACTTTTAAAGAGGTCTCTAATCTGAAGAGATGGAAGCCAGCCCTGCCAAAGATGA 2100  
QY 2099 GATCCATAATATGATGCCAGTGAACCTTATTTGTGAACCATACCCTGCCAATGACTAAGG 2158  
Db 2101 GATCCATAATATGATGCCAGTGAACCTTATTTGTGAACCATACCCTGCCAATGACTAAGG 2160  
QY 2159 AATCAAGAGAGAGAACCAACGTTCTTAAAGTACAGTGAACATATACAATTTGACTGA 2218  
Db 2161 AATCAAGAGAGAGAACCAACGTTCTTAAAGTACAGTGAACATATACAATTTGACTGA 2220  
QY 2219 GTGCAGATTTAGATTTTCATGGGAGCAGCCTCTAATTTAGACAACCTTAAGCAACGTTGATC 2278  
Db 2221 GTGCAGATTTAGATTTTCATGGGAGCAGCCTCTAATTTAGACAACCTTAAGCAACGTTGATC 2280  
QY 2279 GGCTGCTTCTTATCATTTGCTTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTAAT 2338  
Db 2281 GGCTGCTTCTTATCATTTGCTTTTCCATCTAGATCAGTTACAGCCATTGATTCCTTAAT 2340  
QY 2339 GTTTTTCAGTCTTCCAGGATTTGTTAGTTAGCTACTATGTAATTTTCAGGGAAT 2398  
Db 2341 GTTTTTCAGTCTTCCAGGATTTGTTAGTTAGCTACTATGTAATTTTCAGGGAAT 2400  
QY 2399 AGTTAAGCTTTATTCATTCATGCAATACTAAAGAGAAATTAAGTAATGCAATTTTGTG 2458  
Db 2401 AGTTAAGCTTTATTCATTCATGCAATACTAAAGAGAAATTAAGTAATGCAATTTTGTG 2460  
QY 2459 CTGGCTTTGAACAATTTACGAACAATAATGAAGACAATAATCCTGAAGAGAGATTTT 2518  
Db 2461 CTGGCTTTGAACAATTTACGAACAATAATGAAGACAATAATCCTGAAGAGAGATTTT 2520  
QY 2519 AAAAAATGTTTGTCTTCTTCAAAATGAGATTTTTTTGTACCAAGCTTTACCACTTTTC 2578  
Db 2521 AAAAAATGTTTGTCTTCTTCAAAATGAGATTTTTTTGTACCAAGCTTTACCACTTTTC 2580  
QY 2579 AGCCATTTATTAATATGGGAATTTAACTTACTCAAGCAATATGTTGAAGGAGGTGCATA 2638  
Db 2581 AGCCATTTATTAATATGGGAATTTAACTTACTCAAGCAATATGTTGAAGGAGGTGCATA 2640  
QY 2639 TTATCAGGATGCAATTTATGTTGTGTGCCAGTGTGCCAACAATCAATTTCTTAACA 2698  
Db 2641 TTATCAGGATGCAATTTATGTTGTGTGCCAGTGTGCCAACAATCAATTTCTTAACA 2700  
QY 2699 TGAGCTCCAGTTTACCTTAATGTTTCACTGACACAAAGATGAGATTACACCTACAGTGC 2758  
Db 2701 TGAGCTCCAGTTTACCTTAATGTTTCACTGACACAAAGATGAGATTACACCTACAGTGC 2760  
QY 2759 TCTGAGTAGTCACATATTAAGCACTGCACATGAGATATAGATCCGTAGAATTGTGAGA 2818  
Db 2761 TCTGAGTAGTCACATATTAAGCACTGCACATGAGATATAGATCCGTAGAATTGTGAGA 2820  
QY 2819 GTGCACCTCTCTACTTGGGAGGTACAATTGCCATATGATTTCTAGCTGCCATGTTGTTA 2878  
Db 2821 GTGCACCTCTCTACTTGGGAGGTACAATTGCCATATGATTTCTAGCTGCCATGTTGTTA 2880  
QY 2879 GGAATGTGATACTGCTGTTTGGCAAGTCAACAGACCTTGCTCAGAAAGAGCTGTGAGCC 2938  
Db 2881 GGAATGTGATACTGCTGTTTGGCAAGTCAACAGACCTTGCTCAGAAAGAGCTGTGAGCC 2940  
QY 2939 AGTATTCATTTAAGAGAAATTCACACCACTGGCGCGCGCTTGAT 2985  
Db 2941 AGTATTCATTTAAGAGAAATTCACACCACTGGCGCGCGCTTGAT 2987

RESULT 3  
US-10-355-218-3  
; Sequence 3, Application US/10355218  
; Publication No. US20040009558A1  
; GENERAL INFORMATION:  
; APPLICANT: MORAS, DINO  
; APPLICANT: RENAUD, JEAN-PAUL

APPLICANT: GRESCHIK, HOLGER  
APPLICANT: WURIZ, JEAN-MARIE  
TITLE OF INVENTION: THE LIGAND BINDING POCKET PEPTIDE FRAGMENT OF THE  
TITLE OF INVENTION: ESTROGEN-RELATED RECEPTOR 3 (ERR3) AND USES THEREOF  
FILE REFERENCE: 0510-1060  
CURRENT APPLICATION NUMBER: US/10/355,218  
CURRENT FILING DATE: 2003-01-31  
PRIOR APPLICATION NUMBER: 60/352,551  
PRIOR FILING DATE: 2002-01-31  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 1377  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-355-218-3

Query Match 46.1%; Score 1377; DB 16; Length 1377;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 ATGATTCGTAGAACTTGGCTTCTGTAATCTTTTCCCTGCACCTACGAGGAAGACTT 185  
DB 1 ATGATTCGTAGAACTTGGCTTCTGTAATCTTTTCCCTGCACCTACGAGGAAGACTT 60  
QY 186 CTCTGCAGATGTCAACAAGAATCGACACATTTCCAGCTGTTCCTTCATCAAG 245  
DB 61 CTCTGCAGATGTCAACAAGAATCGACACATTTCCAGCTGTTCCTTCATCAAG 120  
QY 246 ACGGAACCTTCCAGCCCGCTCCCTGACGAGACAGCGTCAACCAACAGCCCTGGTGGC 305  
DB 121 ACGGAACCTTCCAGCCCGCTCCCTGACGAGACAGCGTCAACCAACAGCCCTGGTGGC 180  
QY 306 TCTTCAGACGCGAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGACTTGAC 365  
DB 181 TCTTCAGACGCGAGTGGAGCTACAGTTCAACCATGAATGGCCATCAGAACGACTTGAC 240  
QY 366 TCGCCACCTCTCTACCTTCTGCTCTATCTCTGGAGGTAGTGGCCCTGTCAAGAAACTG 425  
DB 241 TCGCCACCTCTCTACCTTCTGCTCTATCTCTGGAGGTAGTGGCCCTGTCAAGAAACTG 300  
QY 426 TATGATGACTGTCCAGACCAATTTGTAAGATCCCCAGACCAAGTGTGAATACATGCTC 485  
DB 301 TATGATGACTGTCCAGACCAATTTGTAAGATCCCCAGACCAAGTGTGAATACATGCTC 360  
QY 486 AACTCGATGCCCCAAGAGACTGTGTGTTAGTGTGTGTGACATCGCTTCTGGTACCACTAT 545  
DB 361 AACTCGATGCCCCAAGAGACTGTGTGTTAGTGTGTGTGACATCGCTTCTGGTACCACTAT 420  
QY 546 GGGTAGCATCATGTGAAGCCTGCAAGGCATTCTTCAAGAGACAATTCAAGGCAATATA 605  
DB 421 GGGTAGCATCATGTGAAGCCTGCAAGGCATTCTTCAAGAGACAATTCAAGGCAATATA 480  
QY 606 GAATACAGCTGCCCTGCCACGAATGATGTAATCACAAAAGCGCAGACGTAAATCTTGC 665  
DB 481 GAATACAGCTGCCCTGCCACGAATGATGTAATCACAAAAGCGCAGACGTAAATCTTGC 540  
QY 666 CAGGCTTGGCGCTTTCATGAAGTGTAAAGTGGGCATGCTGAAAGAAGGGGTGCTCTT 725  
DB 541 CAGGCTTGGCGCTTTCATGAAGTGTAAAGTGGGCATGCTGAAAGAAGGGGTGCTCTT 600  
QY 726 GACAGAGTACGTGGAGGTGGCAGAGTCAAGCGCAGATAGATGGGAGAAACAGCCCA 785  
DB 601 GACAGAGTACGTGGAGGTGGCAGAGTCAAGCGCAGATAGATGGGAGAAACAGCCCA 660  
QY 786 TACCTGAACCTTCAGCTGTTCAGCCGCAAAAAGCCATATATAAGATTTGCTACAT 845  
DB 661 TACCTGAACCTTCAGCTGTTCAGCCGCAAAAAGCCATATATAAGATTTGCTACAT 720  
QY 846 TTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGACCCCTACTGTCCCGACAGT 905  
DB 721 TTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGACCCCTACTGTCCCGACAGT 780

QY 906 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGGTTATCATTT 965  
DB 781 GACATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGGTTATCATTT 840  
QY 966 GGATGGGCGAAGCATATTTCCAGGCTTCTCCAAGCTGTCCCTGGCGGACAGATGAGCCTT 1025  
DB 841 GGATGGGCGAAGCATATTTCCAGGCTTCTCCAAGCTGTCCCTGGCGGACAGATGAGCCTT 900  
QY 1026 CTGCAGAGTGTGGATGGAATTTTGAATCTTGGTGTGCTGATATACCGGCTCTTCAATT 1085  
DB 901 CTGCAGAGTGTGGATGGAATTTTGAATCTTGGTGTGCTGATATACCGGCTCTTCAATT 960  
QY 1086 GAGGATGAAGTGTCTATGCAGACGATTAATTAATGACGGAAGACCAAGTCCAAATTAGCA 1145  
DB 961 GAGGATGAAGTGTCTATGCAGACGATTAATTAATGACGGAAGACCAAGTCCAAATTAGCA 1020  
QY 1146 GGCCCTTCTGATCTAATAATATGCTATCTGACGCTGTAAAGAAATACAAAGCATGAAG 1205  
DB 1021 GGCCCTTCTGATCTAATAATATGCTATCTGACGCTGTAAAGAAATACAAAGCATGAAG 1080  
QY 1206 CTGAAAAAAGAAATTTGTCAACCTCAAGCTATAGCTCTTGTAAATTCAGACTCCATG 1265  
DB 1081 CTGAAAAAAGAAATTTGTCAACCTCAAGCTATAGCTCTTGTAAATTCAGACTCCATG 1140  
QY 1266 CACATAGAAGATGTGAAGCCGTTCAAGAGCTTCAGATGTCTTACATGAAGCGCTGCAG 1325  
DB 1141 CACATAGAAGATGTGAAGCCGTTCAAGAGCTTCAGATGTCTTACATGAAGCGCTGCAG 1200  
QY 1326 GATTATGAAGCTGGCCAGACATGGAAGACCTCTCGTCCAGCTGGCAAGATGCTGATGACA 1385  
DB 1201 GATTATGAAGCTGGCCAGACATGGAAGACCTCTCGTCCAGCTGGCAAGATGCTGATGACA 1260  
QY 1386 CTGCCACTCTGAGGAGACCTCTACCAAGGCCGTGACGATTTCTACACATCAACTA 1445  
DB 1261 CTGCCACTCTGAGGAGACCTCTACCAAGGCCGTGACGATTTCTACACATCAACTA 1320  
QY 1446 GAAGCAAAAGTCCCAATGCACAAACTTTTGGAAATGTGGAGGCCAAGGTCTGA 1502  
DB 1321 GAAGCAAAAGTCCCAATGCACAAACTTTTGGAAATGTGGAGGCCAAGGTCTGA 1377

RESULT 4  
US-10-054-841-1  
Sequence 1, Application US/10054841  
Publication No. US20030119100A1  
GENERAL INFORMATION:  
APPLICANT: Chen, Fang  
TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
TITLE OF INVENTION: RECEPTOR PROTEINS  
FILE REFERENCE: 19999YCA  
CURRENT APPLICATION NUMBER: US/10/054,841  
CURRENT FILING DATE: 2002-01-23  
PRIOR APPLICATION NUMBER: 09/487,379  
PRIOR FILING DATE: 2000-01-18  
PRIOR APPLICATION NUMBER: 09/141,000  
PRIOR FILING DATE: 1998-08-26  
PRIOR APPLICATION NUMBER: 60/078,633  
PRIOR FILING DATE: 1998-03-19  
PRIOR APPLICATION NUMBER: 60/062,902  
PRIOR FILING DATE: 1997-10-21  
PRIOR APPLICATION NUMBER: 60/057,090  
PRIOR FILING DATE: 1997-08-27  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 2807  
TYPE: DNA  
ORGANISM: Human  
US-10-054-841-1

Query Match 24.0%; Score 715; DB 15; Length 2807;  
Best Local Similarity 71.5%; Pred. No. 8.8e-178;  
Matches 956; Conservative 0; Mismatches 375; Indels 6; Gaps 1;

QY	179	AGAGCTTCTCTGCAGAAATGTCAAAACAAGATGCACACATTGATTCCAGCTGTTCGCTT	238
Db	934	AGGGCTGTGAACAAGATGTCTCGAGCAGACAGGACCTGGGCTCCAGCTGGGCTCCTT	993
QY	239	CATCAAGCGGAACCTTCCAGCCAGCCTCCCTGACGGACAGCGTCAACCAACAGCCC	298
Db	994	CATCAAGACTGAGCCGTCCAGCCCGCTCTGGGCATAGATGCCCTCAGCCACCAAGCCC	1053
QY	299	TGGTGGCTCTTCAGAGCCAGTGGAGCTACAGTTCAACCATGAAATGGCCATCAGAACGG	358
Db	1054	CAGTGGCTGTCTCCGACGCCAGCGCGCTTTGGCCTGGCCCTGGGCACCACGCCAACGG	1113
QY	359	ACTTGAATCCCACTCTCTACCCCTTGTCTCTCTATCCTGGAGTAGTGGGCTGTGAG	418
Db	1114	TCTGGAATCGCCACC-----CATGTTTGACAGCGCGCGGTGGAGGCACCCCATGGCG	1167
QY	419	GAAACTGTATGATGACTGTCTCCAGACCATTTGTTGAAGATCCCCAGACCAAGTGAATA	478
Db	1168	CAGAGACTACGAGACTGTGCCAGCGGCATCATGAGAGACTGGGCATCAAGTGGAGTA	1227
QY	479	CATGCTCAACTCGATGCCCCAAGAGACTGTGTTAGTGTGTGTACATCGCTTCTGGTA	538
Db	1228	CATGCTCAACGCCATCCCCAAGCGCTGTGCCTCGTGTGCGGGACATTGCCCTTGCTTA	1287
QY	539	CCACTATGGGGTAGCATCATGTGAAAGCCTGCAGAGCATTTCTTCAAGAGCAATTCAAAG	598
Db	1288	CCAATAAGGGCTGGCCTCTGTGCGAGCCTTGCAAGGCCCTTTTCAAGAGCAATTCAAAG	1347
QY	599	CAATATAGAATACAGCTGCCCTGCCACGAATGAATGTGAATCAAAAGCGCAGACGTAA	658
Db	1348	GAACATTGAGTACAGCTGCCCGGCCACCAACGAGTGCAGATCAACAAACGAGGCCCAA	1407
QY	659	ATCTGCCAGGCTTGCCGCTTTCATGAAGTGTTTAAAGTGGGCATGCTGAAAGAAAGGGT	718
Db	1408	GTCTGCCAGGCTTGCCGCTTTCATGAATGCTCAAGTGGGAGTGTGAAGAAAGTGT	1467
QY	719	GCGTCTTGACAGAGTACGTGGAGTTCGGCAGAATACAGCCAGATAGATGCCGAA	778
Db	1468	GCGCTTGATCGAGTGCCTGGAGGCCGCTCAGAAATACAGCGACGGCTGACTCAGAGAG	1527
QY	779	CAGCCCATACCTGAACCTCAGCTGTTTCAGCCAGCCAAAAAGCCATATAACAAGATTGT	838
Db	1528	CAGCCCATACCTGAAGCTTACAAATTTCTCCACCTGTAAAAAGCCATTGACCAAGATTGT	1587
QY	839	CTCACATTGTTGGTGGCTGAACCGGAGAAGATCTATGCCATGCTGACCTACTGTCCC	898
Db	1588	CTCATACCTACTGTGGCTGAGCCGGACAAGCTCTATGCCATGCTCCCTGCTATGCC	1647
QY	899	CGACAGTACATCAAGCCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGGTGT	958
Db	1648	TGAGGGGACATCAAGCCCTGACCACTCTCTGTGACCTGGCAGACCGAGAGCTTGTGT	1707
QY	959	TATCATTGATGGGCGAAGCATATTCCAGGCTTCTCCACGCTGTCCCTGGCGCAACGAT	1018
Db	1708	CATCATTGCTGGGCGCAAGCAATCCAGGCTTCTCAAGCCTCTCCCTGGGGCAACGAT	1767
QY	1019	GAGCCTTCTGACAGTCTTGGATGGAATTTGATCCTTGGTGTGTAATACCGTCTCT	1078
Db	1768	GAGCCTGCTGACAGTCTTGGATGGAATTTGATCCTTGGTGTGTAATACCGTCTCT	1827
QY	1079	TTCATTGAGGATGAACCTGTCTATGCAGACGATTAATAATGACGAAGACCAAGTCCAA	1138
Db	1828	GCCCTACGACGACAAGCTGTGTACGTGAGGACTACATCATGATGAGGACACTCCCG	1887
QY	1139	ATTAGCAGGCTTCTTGATCTAAATATGCTATCTTCAGCTGTGTAAGAAATACAGAG	1198
Db	1888	CCTGCGGGCTGTGTGAGCTCTACCGGGCATCTGCAGCTGTGTACGCAAGTACAGAA	1947
QY	1199	CATGAAGCTGAAAAAAGAAATTTGTACCCCTCAAAGCTATAGCTTTGCTAATTACA	1258
Db	1948	GCTCAAGGTGAGAAAGAGAGATTGTGTACGCTCAAGGCCCTGGCCCTCGCCAACTCCGA	2007

QY	1259	CTCCATGCACATAGAGAATGTGAAGCCCGTTCAGAAAGCTTCAGGATGTCTTACATGAAGC	1318
Db	2008	TTCCATGTACATCGAGGATCTAGAGGCTGTCCAGAAAGCTGCAGGACCTGCTGCACGAGGC	2067
QY	1319	GCTGCAGGATTATGAAGCTGGCCAGCACAATGGAAGACCTCGTCGAGCTGGCAAGATGCT	1378
Db	2068	ACTGCAGGACTACGAGCTGAGCCAGCGCCATGAGGAGCCTCGAGGAGCAGCGGCAAGCTGCT	2127
QY	1379	GATGACACTGCCACTCCTGAGGCGAGACCTCTACCAAGSCGTGCAGCATTTCTAACACAT	1438
Db	2128	GCTGACACTGCCGCTGCTGTCGGCGCAGACGGCCGCCAAGSCCTGTCAGCACTTCTATAGCGT	2187
QY	1439	CAAACTAGAAAGGCAAAGTCCCAATGCACAAACTTTTTTTGGAAATGTTGGAGGCCAAGGT	1498
Db	2188	CAAACTGCAGGGCAAAAGTCCCAATGCACAAACTTTCTCTGGAGATGCTGAGGCCAAGGC	2247
QY	1499	CTGACTAAAAAGCTCCCT	1515
Db	2248	CTGGGCCAGGGCTGACT	2264

## RESULT 5

```

US-10-101-510-87
; Sequence 87, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 2153
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-87

```

Query Match	23.3%	Score 695;	DB 15;	Length 2153;
Best Local Similarity	69.7%;	Pred. No. 1.4e-172;		
Matches 972; Conservative	0;	Mismatches 415;	Indels 8;	Gaps 2;

OY		108	AATGTTGCCGGTCGCACATGGAATTCCGTAGAACTTTCCTTCTCGAATCTTTTTCCCTG	167
Db		15	AATGCTAAAAACGGACTGATGGA CGTGCCGA ACTTCGATCCCCGACCCTCGGCTAC	74
OY		168	CAC TAC GGA GA AG A C T T C T T G C A G A AT G T C A A C A A G A T G A C A C A T T G A T T C C A G C	227
Db		75	CACAACCA GTA--GGTTGCTGA ACCGAATGTCGTCGAGA GACAGCACCCTGGGCTTAGC	132
OY		228	TGTTGCTCTTCATCAAGACGGAACCTTCCAGCCCAGCCTCCCTGACGGACAGCGTCAAC	287
Db		133	TGCGGCTCTTCATCAAGACGGAACCATTA GCCCATCTCGGGCA TTGATGCCCTCAGC	192
OY		288	CACCACAGCCCTGGTGGCTCTTCAGACGCCAGTGGAGCTACAGTTC AACCATG AATGGC	347
Db		193	CACCACAGCCCCACGGGCTCGTCGAGCCGCA CGGTGGCTTTGGCA TGAGCCCTGGGCA CC	252
OY		348	CATCAGAACGGA CTTGACTCGCCACCCTCTCAACCCTTCTGCTCCTATCCTGGGAGGTAGT	407
Db		253	CACGCCAACGGTCTGGA CTCTCCGCTATGTTGCA GTGC-----GGGGCTGGGAGGC	306
OY		408	GGGCTGTCA GGA A A C T G T A T G A T G A C T G C T C C A G C A C C A T T G T T G A A G A T C C C A G A C C	467
Db		307	AACCCGTGTCCGCAAGAGCTACGAGGACTGTACTAGCGGTATCATGGAGAGACTCGGCATC	366
OY		468	AAGTGTGAATCATGCTCAACTCGATGCCCCA GAGA CTGTGTTAGTGTGTGATC	527
Db		367	AAGTGCAGTACATGCTTAA CGCCATCCCCAAGCGCTGTGCTCTGTGTGCGGGACATT	426



QY 528 GCTTCTGGGTACCATATGGGGTAGCATCATGTGAAGCTTGCAGGCAATCTTTCAAGAG 587  
Db 427 GCTTCTGGCTACCATATGAGTGGCCCTCCTGCGAGGCTTGCAAGCGTTCTTCAAGAGA 486  
QY 588 ACAATTCAAGGCAATATAGATAAGCTGCCCCCTGCCAGATGAATGTGAATCAAAAG 647  
Db 487 ACCATTCAAGAAACATCGAATAACAGCTGCCCTGCCCAACGAGTGTGAGATCACCAAA 546  
QY 648 CGCAGACGTAATCCTGCGAGGCTTGCCGCTTCATGAGTGTAAAGTGGGCGATGCTG 707  
Db 547 CGGAGGCGCAAGTCTGTGAGGCTTGCCGGTTTCAATGAATGCCCTCAAGTGGGATGCTG 606  
QY 708 AAAGAAAGGGGTGCGTCTTGAAGAGTACGTGAGGTGCGGAGAAAGTACAGCGCAGATA 767  
Db 607 AAGGAAGGCGGTGCGCTTGACCGGGTGCAGAGAGGCGCCAGAAAGTACAGAGACGGCTG 666  
QY 768 GATCGGAGAAACAGCCCAATACCTGAACCCCTCAGCTGGTTACGCCCAAAAGCATAT 827  
Db 667 GATTGGAGAAACAGCCCTTACCTGAGCTTACAGATTTCCTCCGCTGCTAAAAAGCATTTG 726  
QY 828 AACAGATTGTCTCACATTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGAC 887  
Db 727 ACTAAGATTGTCTGTATCTACTGTGGCCGAGCCGGAACAAGCTGTACGCTATGCTTCCC 786  
QY 888 CCTACTGCCCCGACAGTGAATCAAAAGCCCTCACTACTGTGTGACTTGGCCGACGA 947  
Db 787 GACGATGTGCTGAAGGGGATCAAGGCCCTGACCACTCTGTGACTTGGCAGATCGG 846  
QY 948 GAGTGGTGTATCATTTGAGTGGGCGAAGCATATTCAGGCTTCTCCACGCTGCTCCG 1007  
Db 847 GAGCTTGTGTCTCTCATATGCTGGGCCAAGCACATCCCAAGTTCTCCAACTGACACTC 906  
QY 1008 GCGGACCAGATGAGCCTTCTGAGAGTGTGCTTGATGGAATTTTGATCCTTGGTGTG 1067  
Db 907 GGGGACCAAGATGAGCCTGCTGAGAGTGCCTGATGAGATTCCTCATCCTGGGCATGCTG 966  
QY 1068 TACCGGTCTCTTTCATTGAGATGAATCTGTCTATGAGAGATTTATATATGAGCAA 1127  
Db 967 TACCGGTCTCTTCTTATGATGAGAGTGTGCAATACGCGGAGACTATATCATGATGAG 1026  
QY 1128 GACCACTCCAAATTTAGCAGGCTTCTGTGATCTAATTAATGCTATCTGACGCTGTAAG 1187  
Db 1027 GAACACTCTGCTGCTGAGGCTGCTGGAAGCTTTACCGAGCATCTTGACGCTGTACGC 1086  
QY 1188 AAATACAAGAGCATGAAGCTGGAAGAAAGAAATTTGTCAACCTCAAAAGCTATAGCTTT 1247  
Db 1087 AGGTACAAGAAAGCTCAAGGTGAGAGAAAGATTGTGATGCTCAAAAGCCTGGCCTT 1146  
QY 1248 GCTAATTCAGACTCCATGACATAGAGATGTGTAAGCCGTTCAAGGCTTCAGATGTC 1307  
Db 1147 GCCAATCAGATTCAATGATGAGAAAGCTGAGGCTGTGAGAGGCTTCAGGACCTG 1206  
QY 1308 TTACATGAAGCGCTGACGAGTATGAAAGCTGGCCAGCACATGGAAGACCTCGTGAGCT 1367  
Db 1207 CTGCAATGAGGCGCTGCAAGACTATGAGCTGAGCCAGCCCATGAGAGGCCACGAGGGCG 1266  
QY 1368 GGCAGATGCTGATGACACTGCACTCTGAGGCAAGACCTTACCAAGGCCGTGACGAT 1427  
Db 1267 GGCAGCTGCTGTTGACACTGCCCTGCTGCGGCAAGCGCAAGCCCAAGCCGTCCAGCAC 1326  
QY 1428 TTCTACAACATCAAACTAGAGGCAAGTCCCAATGACAAACTTTTGTGAAATGTG 1487  
Db 1327 TTCTACAGTGTGAAACTGAGGCGCAAGGTGCCCATGACAAACTCTTCTGAGATGCTG 1386  
QY 1488 GAGGCCAAGGTCTGA 1502  
Db 1387 GAGGCCAAGGTCTGA 1401

RESULT 6  
US-10-101-510-413  
; Sequence 413, Application US/10101510

; Publication No. US20030148295A1  
; GENERAL INFORMATION:  
; APPLICANT: WAN, JACKSON  
; APPLICANT: WANG, YIXIN  
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
; FILE REFERENCE: 15117.0012  
; CURRENT APPLICATION NUMBER: US/10/101,510  
; PRIOR FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/276,947  
; NUMBER OF SEQ ID NOS: 805  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 413  
; LENGTH: 2153  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-101-510-413  
  
Query Match 23.3%; Score 695; DB 15; length 2153;  
Best Local Similarity 69.7%; Pred. No. 1.4e-172;  
Matches 972; Conservative 0; Mismatches 415; Indels 8; Gaps 2;  
  
QY 108 AATGTTGCCGTCGACATGATTCGGTAGAACTTGCTTCTGCAATCTTTTCCCTG 167  
Db 15 AATGCTAAACGGGACTGATGAGCTGTCCGAATCTGCAATCCCGAACCCCTCGGCTAC 74  
QY 168 CACTACGAGGAAGAGCTTCTGCAAGATGTCAACAAGATCGACACATGATTCAGC 227  
Db 75 CACAACCAAGTA--GGTTGCTGAACCGAATGTGTCGCGAAGACAGGCACTGGGCTCTAGC 132  
QY 228 TGTTCGTCCTTCATCAAGACGGAACCTTCCAGGCCCGCTCCCTGACGAGACGCTCAAC 287  
Db 133 TCGGGCTCTTCATCAAGACGGAAGCCATCTAGCCCATCTCGGGCAATGATGCCCTGAGC 192  
QY 288 CACCAACAGCCCTGTGCTCTTCAAGCGCCAGTGGAGCTACAGTTCAACCATGATGAGC 347  
Db 193 CACCAACAGCCCGAGCGGCTGTGAGACGCGAGCGGCTGTTGGCATGGCCCTGGGCAAC 252  
QY 348 CATCAGAAGGACTTGACTGCGCACTCTCTACCTTCTGCTCTATCCTGAGAGTACT 407  
Db 253 CAGGCCAACGCTGTGACTCTCCGCTATGTTCGCAAGTGC-----GGGCTGGGAGGC 306  
QY 408 GGGCCTGTCAGGAACCTGTATGATGACTGCTCCAGCACCATTTGTAAGATCCCGACACC 467  
Db 307 AACCGTGTGCAAGAGCTTCAAGAGACTGTACTAGCGGTATCATGAGAGACTCGGCCATC 366  
QY 468 AAGTGTGATACATGCTCAACTCGATGCCCAAGAGACTGTGTTAAGTGTGTGATCATC 527  
Db 367 AAGTGTGAGTACATGCTTAAAGCCATCCCAAGCGCTGTGCTGCTGTGCGGGAGCATTT 426  
QY 528 GCTTCTGGGTACCACTATGAGGTAAGCATCATGTGAAGCCTGCAAGGCAATTTCTCAAGAG 587  
Db 427 GCTTCTGGCTACCACTATGAGGTGGCTCTGCGAGGCTTGCAAGGCGTTCCTCAAGAGA 486  
QY 588 ACAATTCAAGGCAATATAGATAAGCTGCCCCCTGCCAGATGAATGTGAATCACAAG 647  
Db 487 ACCATTCAAGAAACATCGAATAACAGCTGCCCTGCCCAACGAGTGTGAGATCACCAAA 546  
QY 648 CGCAGACGTAATCCTGCGAGGCTTGCCGCTTCATGAGTGTAAAGTGGGCGATGCTG 707  
Db 547 CGGAGGCGCAAGTCTGTGAGGCTTGCCGGTTTCAATGAATGCCCTCAAGTGGGATGCTG 606  
QY 708 AAAGAAAGGGGTGCGTCTTGAAGAGTACGTGAGGTGCGGAGAAAGTACAGCGCAGATA 767  
Db 607 AAGGAAGGCGGTGCGCTTGACCGGGTGCAGAGAGGCGCCAGAAAGTACAGAGACGGCTG 666  
QY 768 GATCGGAGAAACAGCCCATACCTGAACCCCTCAGCTGTGTTACGCCCAAAAGCCATAT 827  
Db 667 GATTGGAGAAACAGCCCTTACCTGAGCTTACAGATTTCCCGCTGCTTAAAAAGCCATTTG 726  
QY 828 AACAGATTGTCTCACATTTGTTGGTGGCTGAACCGGAGAGATCTATGCCATGCTGAC 887  
Db 727 ACTAAGATTGTCTGTATCTACTGTGTGGCCGAGCGGAGCAAGCTGTACGCTATGCTCCC 786

QY	888	CCTACTGTCCCCGACAGTGCATCAAAAGCCCTCACTACATCTGTGTGACTTGGCCGACCGA	947
Db	787	GACGATGTGCTGAAGGGGATATCAAGGCCCTGACCACTCTCTGTGACTTGGCAGATCGG	846
QY	948	GAGTGTGTTATCATTTGATGGCGAAGCATATTTCCAGGCTTCTCCAGCGTGTCCCTG	1007
Db	847	GAGCTGTGTTCTCATTTAGCTGGGCCAAGCACATCCAGGTTTCTCCAACTGCACACTC	906
QY	1008	GCGGACCAGATGAGCCCTTCTGCAGAGTGCCTTGATGGAATTTTGATCCCTGTGTGCTA	1067
Db	907	GGGGACCAGATGAGCCTGTCTGCAGAGTGCCTTGATGGAATTCCTCATCCTGGCATCGTG	966
QY	1068	TACCGGTCTTTTCATTTGAGGATGAACCTTGCTATGCAGACGATTAATATGACGAA	1127
Db	967	TACCGTCTGCTTCCCTATGATGACAAAGCTGGCATACGGGAGGAACTATATCATGATGAG	1026
QY	1128	GACCAGTCCAAATTAGCAGGCCCTTCTTGATCTAAATTAATGCTATCCTGCAGCTGTAAAG	1187
Db	1027	GAACACTCTGCGCTGTGGGGCTGCTGAGCTTTACCGAGCATCTTGACGCTCGTAAGC	1086
QY	1188	AAATACAAGAGCATGAAGCTGGA AAAAGAAGATTTGTCACCCTCAAAGCTATAGTCTT	1247
Db	1087	AGGTACAAGAAGCTCAAGGTGAGAAGAAGAGATTGTGTATGCTCAAAAGCCCTGGCCCTT	1146
QY	1248	GCTAATTCAGACTCCATGCAATAGAAGATGTTGAAGCCGTTCAAGAGCTTCAGAGTGTG	1307
Db	1147	GCCAACTCAGATTCAATGTATCATCGAAGAACCTGGAGGCTGTGCAGAGCTTCAGGACTG	1206
QY	1308	TTACATGAAGCGCTGCAGATTAATGAAGCTGGCCAGCACATGGAAGACCCCTGTCGAGCT	1367
Db	1207	CTGCATGAGGCGCTGCAGACTATAGAGCTGAGCCAGCGCCATGAGGACCAACGAGGGCG	1266
QY	1368	GGCAAGATGCTGATGACACTGCCACTCCTGAGGCGAGACTCTAACCAAGGCCGTGCAGCAT	1427
Db	1267	GGCAAGCTGCTGTTGACACTGCCCTCTGCTGGCGCAGCGGCAAGCCCGTCCAGCAC	1326
QY	1428	TTCTACACATCAAACTAGAAGGCAAAAGTCCCAATGCACAAACCTTTTGGAAATGTTG	1487
Db	1327	TTCTACAGTGTGAACCTGCAGGGCAAGGTGCCCATGCACAAACCTTCTTGAGATGCTG	1386
QY	1488	GAGGCCAAGGTTCTGA 1502	
Db	1387	GAGGCCAAGGTTCTGA 1401	

```

RESULT 7
US-10-054-841-8/c
; Sequence 8, Application US/10054841
; Publication No. US20030119100A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; TITLE OF INVENTION: RECEPTOR PROTEINS
; FILE REFERENCE: 19999YCA
; CURRENT APPLICATION NUMBER: US/10/054,841
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: 09/487,379
; PRIOR FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/141,000
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/078,633
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/062,902
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/057,090
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human

```

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ., (622)
; OTHER INFORMATION: n = A, T, C or G
US-10-054-841-8

```

Query Match	14.3%	Score 428;	DB 15;	Length 622;
Best Local Similarity	93.9%;	Pred. No. 2.9e-102;		
Matches 495; Conservative	0;	Mismatches 27;	Indels 5;	Gaps 5;

QY	25	CGGCTTGCTAC	TAGGAGAACATTTGTGTAA	TTGCCACTGTGCTCTGTCAAGAACTTTG	84
Db	546	CAGTTGATCCCC	NANAGGAGACATTTGTAA	TTGCCCTGT - CTCTGTCAAGG - AACTTTG	489
QY	85	ATTATAGCTGGGG	TGCACAAATPA	TGTTGCCGGTGCACATGATTC - GGTAAGACTT	143
Db	488	ATTAATAGCTGGGG	TGCNCAAAATPA	TGTTTCCGGTGCACATGATTCGGAGAACTT	429
QY	144	TGCCTTCTGAATCTTTT	TCCCTGCATA	C - GAGGAAGAGCTTCTCTGCAGAAATG - CAA	201
Db	428	TGCCTTCTGAATCTTTT	TCCCTGCANACGAGGAAGAGCTTCTCTGCAGAAATG	CA	369
QY	202	ACAAAGATGCACACAT	TGATTCAGCTGTTCCTTCATCAAGACGAACTTCCAGCC		261
Db	368	ACAAAGATGCACACAT	TGATTCAGCTGTTCCTTCATCAAGACGAACTTCCAGCC		309
QY	262	CAGCCTCCCTGACG	GACAGCGTCAACCA	CAGCCCTGTGGCTCTTCAGACGCCAGTG	321
Db	308	CAGCCTCCCTGACG	GACAGCGTCAACCA	CAGCCCTGTGGCTCTTCAGACGCCAGTG	249
QY	322	GGAGCTACAGTTCA	ACCATTGAATGGCCATCAGAACGACTTGACTGGCACTCTCTACC		381
Db	248	GGAGCTACAGTTCA	ACCATTGAATGGCCATCAGAACGACTTGACTGGCCACTCTCTACC		189
QY	382	CTTCTGCTCCTATCCT	TGGAGGTAGTGGGCTGTCAAGAACTGTATGATGACTGTCCA		441
Db	188	CTTCTGCTCCTATCCT	TGGAGGTAGTGGGCTGTCAAGAACTGTATGATGACTGTCCA		129
QY	442	GCAACCATTTGTTGA	AGATCCCCAGACCAAGTGTGAATACATGCTCAACTGCGTCCAGGA		501
Db	128	GCAACCATTTGTTGA	AGATCCCCAGACCAAGTGTGAATACATGCTCAACTGCGTCCAGGA		69
QY	502	GACTGTGTTAGTGTGT	GTGACATGCTTCTGGGTACCACATATGGG		548
Db	68	GACTGTGTTAGTGTGT	GTGACATCNCCTTCNGGGTACCACATTTAGG		22

```

RESULT 8
US-09-962-436-268
; Sequence 268, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatur
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 268
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-268

```

Query Match	14.2%;	Score 424.4;	DB 9;	length 2218;
Best Local Similarity	64.7%;	Pred. No. 6.4e-101;		
Matches 682; Conservative	0;	Mismatches 336;	Indels 36;	Gaps 2;

QY	481	TGCTCAACTCGATGCCCAAGAGACTGTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACC	540
Db	386	TGCTCAGCTCCCTGCCCAAGCGCCTCTGCCTGGTCTGTGGGACGTGGCTCCGGCTACC	445
QY	541	ACTATGGGGTAGCATCATGTGAAGCCCTGCAAGCAATCTTCAAGAGCAATTCAGGCA	600
Db	446	ACTATGGTGTGGCATCCTGTGAGGCTGCAAGCCTTCTTCAAGAGACCATCCAGGGA	505
QY	601	ATATAGATACAGCTGCCCTGCCACGAATGAATGTGAATCACAAAGCGACGTAAT	660
Db	506	GCATCGAGTACAGCTGTCCGGCCTCCACAGAGTGTAGATCACCAAGCGGAGACGCAAG	565
QY	661	CCTGCCAGGCTTCCCGCTTCATGAAGTGTTTAAAGTGGGCATGCTGAAGAAGGGGTGC	720
Db	566	CCTGCCAGGCTTCCCGCTTCACCAAGTGCCTGCCGGTGGGCATGCTCAAGAGGGAATGC	625
QY	721	GTCCTGACAGAGTACGTGGAGGTGCGCAGAGTACAAGCGCAGATGATGCGGAGAACA	780
Db	626	GCCTGAGCCGCTCCCGGGGTGGCGCGCAGAAAGTACAAGCGCGCGGAGGTGAACCCAC	685
QY	781	GCCCATACCTGAACCTCAGCTGGTTACGC-----AGCCA	816
Db	686	TGCCCTTCCCGGGCCCTTCCCTGCTGGGCCCCGTGGCAGTGCCTGAGAGCCCCCGAAGA	745
QY	817	AAAAGCCATATACAGATTTCTCACATTTGTGTGGCTGAACCCGAGAGAATCTATG	876
Db	746	CAGCCCCAGTGAATGCACTGTGTCTCATCTGCTGTGTGTGAGCCTGAGAACCTTATG	805
QY	877	CCATGCCCTGACCCCTACTGTGTCCCGACAGTGCATCAAAAGCCCTACTACATCTGTGACT	936
Db	806	CCATGCCCTGACCCCGCAGGCCCTGATGGGCACCTCCAGCCGTGGCTACCTCTGTGAAC	865
QY	937	TGGCCGACCGAGAGTTGTGTATTATCATTTGATGGCGAAGCATATTCAGGCTTCTCCA	996
Db	866	TCTTTGACCGAGAGATTGTGTCAACCATCAGCTGGGCCAAGACATCCAGGCTTCTCAT	925
QY	997	CGCTGTCCCTGGCGGACCAAGATGAGCCTTCTGCAGAGTCTTGATGGAATTTGATCC	1056
Db	926	CGCTGTCCCTGTCTGACCAAGATGTCAAGTACTGCAGAGCCTGTGATGAGGTGCTGTGC	985
QY	1057	TTGCTGTCTATACCGGTCTCTTCAATTGAGGATGAATGTTCTATGCAGACGATTA	1116
Db	986	TGGGTGTGCCCAGCGCTCACTGACCACGAGATGAGCTGGCCTTGCCTGAGGACTTAG	1045
QY	1117	TAAATGACGAGACCAAGTCCAAATTAGCAGGCCCTTCTGATCTAAATAATGTATCTGC	1176
Db	1046	TCCGTGATGAAGAGGGGCGACGGCAGCTGCGCTGGGGGAACTGGGGCTGCCCTGCTGC	1105
QY	1177	AGCTGTAAAGAATAACAGAGCATGAAGCTGGAAGAAAGAAATTTGTCAACCTCAAG	1236
Db	1106	AACTAGTCCGGCGGCTGACAGGCCCTGCGGCTGAGCGAGAGGAGTATGTTCTACTAAAG	1165
QY	1237	CTATAGCTCTTGCTAATTCAAGCTCCATGCACATAGAGAGTGTGAAGCCGTTCAAGC	1296
Db	1166	CCTTGGCCCTTGCCAAATTCAGACTCTGTGCACATCGAAGATGCCGAGCTGTGGAGCAGC	1225
QY	1297	TTCAAGATGTCTTACATGAAGCGCTGCAGGATTATGAAGCTGGCCAG-----C	1344
Db	1226	TGCGAAGAGCTCTGCACAGAGCCCTGTGAGTATGAAGCGCGCGGCTGGCCCCGAG	1285
QY	1345	ACATGGAAGACCCCTGCTGAGCTGCGCAAGATGCTGATGACACTCCACTCGTAGGCGA	1404
Db	1286	GGGTGTGTAGCGGCGCGCGGCGGCGAGGCTGTGCTCAGCGTACCGCTCCTCCGCGAGA	1345
QY	1405	CCTTACCAAGCGCGTGCAGCATTTCTACAACATCAAACTAAGCAAGCAAGTCCCAATGC	1464
Db	1346	CAGCGGCGAAAGTGTCTGGCCCATTTCTATGGGGTGAAGCTGGAAGGCAAGTGCCTATGC	1405
QY	1465	ACAACTTTTTTTGGAAATGTTGGAGGCCAAGGT	1498
Db	1406	ACAACTGTCTTGTGAAGTGTCTGAGGCCATGAT	1439

```

RESULT 9
US-10-776-827-4
; Sequence 4, Application US/10776827
; Publication No. US20040132086A1
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Kathryn
; APPLICANT: Richer, Jennifer
; TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Related
; TITLE OF INVENTION: Thereto
; FILE REFERENCE: 2848-39
; CURRENT APPLICATION NUMBER: US/10/776,827
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US/09/814,915
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/214,870
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-776-827-4

```

Query Match	14.28;	Score 424.4;	DB 17;	Length 2218;
Best Local Similarity	64.78;	Pred. No. 6.4e-101;		
Matches 682;	Conservative 0;	Mismatches 336;	Indels 36;	Gaps 2;

QY	481	TGCTCAACTCGANGCCCAAGAGACTGTGTGTTAGTGTGTGTGACATCGCTTCTGGGTACC	540
Db	386	TGCTCAGCTCCCTGCCCAAGCGCTCTGCTGTGTGTGGGACGTGGCTTCGGCTACC	445
QY	541	ACTATGGGGTAGCATCATGTGAAGCCTGCAAGCATTCTTCAAGAGACAATTCAAGGCA	600
Db	446	ACTATGCTGTGGCATCTCTGTAGAGCCTGCAAGCCTTCTTCAAGAGACCATCCAGGGGA	505
QY	601	ATATAGATACAGCTGCGCTGCCACGAATGATGTGAATTCACAAAGCGACAGCTAAT	660
Db	506	GCATCGAGTACAGCTGTCTGGGCTCCAAACGAGTGTGAGATCACCAAGCGGAGACGCAAG	565
QY	661	CTGCCAGGCTTGCCGCTTCATGAAGTGTTTAAAGTGGGCATGCTGAAGAAGGGTGC	720
Db	566	CTTGCCAGGCTTGCCGCTTCACCAAGTGCCGTGGGGTGGCATGTCTCAAGGAGGAGTGC	625
QY	721	GTCTTGACAGAGTACGTGAGGTCGGCAGAGTACAAAGCGCAGATAGATGCCGAGAACA	780
Db	626	GCTTGAGCCGCTCCGGGTGGGGCGGCAAGATACAAAGCGGCGGAGGTGAACCCAC	685
QY	781	GCCCATACCTGAACCCCTCAGCTGTTAGCC-----AGCCA	816
Db	686	TGCCCTTCCCGGCCCCCTTCCCTGTGGGCCCTGCGAGTCGCTGAGGCCCGGAAGA	745
QY	817	AAAGCCATATACACAGATTGTCTCACATTTGTTGGTGGCTGAACCGAGAAATCTATG	876
Db	746	CAGCCCAAGTGAATGCACTGTGTCTCATCTGTGTGTGTGAGCCTGAGAAAGCTCTATG	805
QY	877	CCATGCTGACCCCTACTGTCTCCCGACAGTGAACATCAAAAGCCCTCACTACACTGTGTGACT	936
Db	806	CCATGCTGACCCCGCAGGCCCTGATGGCACTCCAGCGGTGGCTACCTCTGTGACC	865
QY	937	TGGCCGACCGAGAGTTGGTGTATCATTTGATGGGCGCAAGCATATTCCAGGCTTCTCA	996
Db	866	TCTTGAACGAGAGATTGTGTCAACATCAGCTGGGCCAAGACATCCAGGCTTCTCAT	925
QY	997	CGCTGTCCCTGGCGGACCAAGATGAGCCTTCTGACAGTGTGATGGAATTTGATCC	1056
Db	926	CGCTGTGCTGTCTGACCAAGATGTCACTACTGACAGCGCTGTGATGAGTGTGTGC	985
QY	1057	TTGGTGTGATATACCGTCTCTTTCATTTGAGATGAATGTCTATGACAGCAATTATA	1116
Db	986	TGGGTGTGCCCGGCTCACATGCACTGCAAGATGAGTGGCCTTGTGCTGAGGACTTAG	1045







Db 458 ACTATGTTGGCATCCTGTGAGCGCTGCAAGCCTTCTTCAAGAGGACCATCCAGGGGA 517  
QY 601 ATATAGAATACAGCTGCCCTGCCACGAATGATGTGAATCACAAGCGCAGACGTAAT 660  
Db 518 GCATCGAGTACAGCTGTCCGGCCTCCACAGACTGTGATCATCCAAGCGGAGACGCAAGG 577  
QY 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGTGAAGAAGGGGTGC 720  
Db 578 CCTGCCAGGCTTGCCGCTTCACCAAGTCTGCGGGTGGGCATGTCAAGAGGAGTGC 637  
QY 721 GTCTTGACAGAGTACGTGAGGTGCGCAGAGTACAAGCGCAGATGATGCGGAGAAC 780  
Db 638 GCCTGACCGCGTCCGGGGTGGCGGCAAGTACAAGCGCGCGGAGGTGACCCAC 697  
QY 781 GCCCATACCTGGAACCTCAGCTGGTTCAAGC-----AG 813  
Db 698 TGCCCTTCCCGGGCCCTTCCCTGTGGGCCCTTGACAGTCGTGAGAGCCCGGAGAGA 757  
QY 814 CCAAAAAGCCATATAACAAGATGTCTCAATTGTTGGTGGCTGAACCGGAGAGATCT 873  
Db 758 CAGACGCCCAAGTGAATGCACTGTGTCTCATCTGTGTGTGAGCTGAGAAGCTCT 817  
QY 874 ATGCCATGCTTGAACCTTACTGTCCCGCAGATGACATCAAGCCCTCACTACACTGTGTG 933  
Db 818 ATGCCATGCTTGAACCTTACTGTCCCGCAGAGCTGTGAGGCACTCCACGCCGTGCTA 877  
QY 934 ACTTGCCGACCGAGAGTGTGTGTTATCACTTGGATGGGCGGAGCATATCCAGGCTTCT 993  
Db 878 ACCCTTTGACCGAGAGATTTGTGTCACTCACTGAGCTGGGCGCAAGACATCCAGGCTTCT 937  
QY 994 CCACGCTGTCCCTGGCGGACCGAGTGAACCTTCTGACAGAGTGTGGATGGAATTTTGA 1053  
Db 938 CATCGCTGTGCTGTCTGACCAAGTGTCACTGACAGAGCGTGTGATGAGTGTGCTGG 997  
QY 1054 TCCCTGTGTGTGTATACCGGCTCTCTTTTCACTTGGAGATGAACCTTGTATGACAGCAT 1113  
Db 998 TGCTGGGTGTGGCCCGCAGCGCTCACTGCCACTGACAGATGAGTGTGCTTGTGAGACT 1057  
QY 1114 ATATATGAGCAGAGACAGCAGTCCAAATTAGCAGGCTTCTGTATTAATATGCTATCC 1173  
Db 1058 TAGTCTTGATGAAGAGGGGCGGACAGCTGGCCTGGGGGAACTGGGGGCTGCCCTGC 1117  
QY 1174 TGCAGCTGTTAAGAATACAGAGCATGAAGCTGGAAGAAAGAAATTTGTCAACCTTCA 1233  
Db 1118 TGCACTAGTGGCGCGCGCTGCAAGCGCTTCCGGCTGAGCGCAGAGAGATATGTTCTTAA 1177  
QY 1234 AAGCTATAGCTCTTGTATTAATCACTCCATGCACATAGAAGATGTTGAAGCCGTTAGA 1293  
Db 1178 AGGCTTGGCCCTTGCCAATTCAGACTCTGTGCACATGAAAGATGCCGAGGCTGTGAGC 1237  
QY 1294 AGCTTCAAGATGTCTTACATGAAGCGCTGCAGGATTATGAAGCTGGCCAG----- 1343  
Db 1238 AGCTGCGAGAGAGCTCTGCACGAGGCCCTGTGTGAGTATGAAGCGCGGCTGGCCCG 1297  
QY 1344 --CACATGGAAGACCCCTGTCGAGCTGGCAAGATGCTGATGACACTGCCACTCTGAGGC 1401  
Db 1298 GAGGGGTGTGAGCGCGCGCGGCGGAGAGCTGTCTCAAGCTACCGCTCTCCGCC 1357  
QY 1402 AGACCTTACCAAGCGCGTGCAGATTCTTCAACAACATCAAACTAGAAGGCAAGTCCCA 1461  
Db 1358 AGACAGCGGCGCAAGTGTGCGCCATTCTATGGGTGAAGCTGAGGCGCAAGGTGCCA 1417  
QY 1462 TGCACAAACTTTTTTGGAAATGTTGAGGCCAAGGT 1498  
Db 1418 TGCACAAAGCTGTCTTGGAGATGCTCGAGGCCATGAT 1454

RESULT 13  
US-10-342-887-1041  
; Sequence 1041, Application US/10342887  
; Publication No. US20040058340A1  
; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Christopher J.  
; APPLICANT: Van 't Veer, Laura Johanna  
; APPLICANT: Van de Vijver, Marc J.  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-188-999  
; CURRENT APPLICATION NUMBER: US/10/342,887  
; PRIOR FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: 60/298,918  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/380,710  
; PRIOR FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: 10/172,118  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 1041  
; LENGTH: 2402  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-342-887-1041

Query Match 13.5%; Score 401.6; DB 13; Length 2402;  
Best Local Similarity 64.5%; Pred. No. 7.4e-95;  
Matches 684; Conservative 0; Mismatches 334; Indels 42; Gaps 4;

QY 481 TGCTCACTCGATGCCCAAGAGACTGTGTTAGTGTGTGTGACATCGCTTCTGGTACC 540  
Db 563 TGCTAGCTCCCTGCCCAAGCGCTCTGCTGTGTGTGGGACGTGGCTCCGGCTACC 622  
QY 541 ACTATGGGTAGCATCATGTGAAGCTGCAAGGCATCTTCAAGAGCAATTCAGGCA 600  
Db 623 ACTATGGGTAGCATCTGTGAGGCTTCAAGGCCTTCTCAAGAGCAATTCAGGCA 682  
QY 601 ATATAGAATACAGCTGCCCTGCCACGAATGAATGTAATCAAAAGCGCAGACGTAAT 660  
Db 683 GCATCAAGTACAGCTGTCCGGCTTCAACGAGTGTGATCAACGAGCGGAGCAAGG 742  
QY 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCAATGCTGAAGAGGGGTGC 720  
Db 743 CTTGCCAGGCTTGCCGCTTCAACGAGTGTGCTGCGGTGGGCAATGCTGAAGAGGGTGC 802  
QY 721 GTCTTGACAGAGTACGTGAGGTGCGCAGAGTACAAGCGCAGATGATGCGAGAGACA 780  
Db 803 GCTTGACCGCGTCCGGGTGGCGGAGAGTACAAGCGCGGCGGAGGTGACCCAC 862  
QY 781 GCCCATACCTGAACCTCAGCTGGTTCAAGC-----AG 813  
Db 863 TGCCCTTCCCGGCGCCCTTCCCTGTGAGGCGCCCTGCAATGCTGAGAGGCCCGGAGAGA 922  
QY 814 CCAAAAAGCCATATAACAAGATTTGTCTACATTTGTGTGGCTGAACCGGAGAGATCT 873  
Db 923 CAGGACCCCAAGTGAATGCACTGTGTCTCATCTGTGTGTGAGCCTGAGAGACTCT 982  
QY 874 ATGCCATGCTGACCCCTACTGTCCCGCAGAGTGAATCAAAAGCCCTCACTACACTGTGTG 933  
Db 983 ATGCCATGCTGACCCCGCAGAGCCCTGATGGGCACTCCAGCGGTGCTACTCTGTG 1042  
QY 934 ACTTGCCGACCGAGAGTGTGTGTTATCATTTGATGGGCGAAGCATATTCAGGCTTCT 993  
Db 1043 ACTCTTTGACCGAGAGATGTGTACCATCAGCTGGCGCAAGCATCCAGGCTTCT 1102  
QY 994 CCACGCTGTCCCTGGCGGACAGATGAGCCTTCTGACAGTGTGATGGAATTTTGA 1053  
Db 1103 CATGCTGTGCTGTGTGACCAAGATGTCACTGAGAGCGTGTGATGAGGTGCTGG 1162  
QY 1054 TCCCTGTGTGTATACCGGTCTCTTCACTTGAAGTGAATCTTGTATGACAGCAT 1113  
Db 1163 TGCTGGGTGTGGCGGCGCTCACTGCCACTGCAAGATGAGCTGGCCTTGTGAGGACT 1222



```
QY 1114 ATATATGACGAAGACAGTCCAATTAGCAGCCTTCTTGATCTAAATAATGCTATCC 1173
    |||||
Db 1223 TAGTCTTGATGAAGAGGGGGCAGCGCAGCTGGCTGGGGAACTGGGGCTGCCCTGC 1282
QY 1174 TGCAGCTGTTAAAGAAATACAGAGCATGAGCTGAGAAAAGAGAATTTGTCACCCTCA 1233
    |||||
Db 1283 TGCACCTAGTGGCGCGCTGAGGCCCTCGCGCTGAGCGAGAGAGTATGTTCTACTAA 1342
QY 1234 AAGCTATAGCTCTTGTAAATTCAGACTCCATGCACATGAGAAT--GTTGAAGCGGTTC 1291
    |||||
Db 1343 AGGCTTGGCCCTTGCCAATTCAGACTCTGTGCACATGAGAATGAGCCGAGCTGTGA 1402
QY 1292 GAAGCTTC-AGGATGTCTTACATGAAGCGCTGACAGATTATGAAGCTGGCCAG----- 1343
    |||||
Db 1403 GCAGCTGCAGAAGACTCCTGCACGAGCCCTGTGAGTATGAAGCCGCGGCTGGCC 1462
QY 1344 -----CACATGGAAGACCTCGTTCAGCTGCAAGATGCTGATGACACTGCCACTCCTGA 1398
    |||||
Db 1463 CCGAGAGGGGTGCTGAGCGGGCGGGCGGCGCAGCTGCTGCACGCTACCGCTCCTCC 1522
QY 1399 GGCAGACCTCTACCAAGCGCGCTGCAGCATTTCTACAACATCAAACTAGAAGGCAAGTCC 1458
    |||||
Db 1523 GCCAGACAGCGGGGCAAAAGTCTGCCCATTTCTATGGGGTGAAGCTGGAGGGCAAGTGC 1582
QY 1459 CAATGCACAAACTTTTGTGGAATGTTGAGGCCCAAGT 1498
Db 1583 CCATGCACAAAGCTGTTCTTGAGATGCTCGAGCCATGAT 1622
```

```
RESULT 14
US-10-172-118-1041
; Sequence 1041, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1041
; LENGTH: 2402
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_004451
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1041
```

Query Match 13.5%; Score 401.6; DB 13; Length 2402;  
Best Local Similarity 64.5%; Pred. No. 7.4e-95;  
Matches 684; Conservative 0; Mismatches 334; Indels 42; Gaps 4;

```
QY 481 TGCTCACTCGATGCCCAAGAGACTGTGTTAGTGTGTGATCATCGCTTCTGGGTACC 540
    |||||
Db 563 TGCTCAGCTCCCTGCCCAAGCGCCTGTGCTGTGTGGGAGCTGGCTCCGGCTACC 622
QY 541 ACTATGGGTAGCATCATGTGAAGCCTGCAAGCATTTCTCAAGAGACATTCAGGCA 600
    |||||
Db 623 ACTATGTTGTGATCTCTGTGAGGCTGCAAGCCTTCTCAAGAGACATTCAGGGA 682
QY 601 ATATGAATACAGTGCCTGCGCAGAAATGAATGAAATCACAAGCGCAGACGTAAAT 660
    |||||
Db 683 GCATCGAGTACAGCTGTCGGCCTCCAACGAGTGTGAGATCACCAGCGGAGACGCAAGG 742
```

```
QY 661 CCTGCCAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGCATGCTGAAGAAGGGTGC 720
    |||||
Db 743 CCTGCCAGGCTTGCCGCTTCACCAAGTCCCTGCGGGTGGGCATGCTCAAGAGAGAGTGC 802
QY 721 GTCTTGACAGATGACGTGAGAGTCCGCGAGAGTACAAAGCGCAGATGATGCCGAACA 780
    |||||
Db 803 GCCTGAGACCGGTCCGGGTGGCGCGCAGAAATCAAGCGCGCGGAGGTGACCCAC 862
QY 781 GCCCATACCTGAACCTCAGCTGTTACGC-----AG 813
    |||||
Db 863 TGCCCTTCCGGGCCCTTCCCTGTGAGCCCTGCGAGTCCGTGAGGCCCGCGAAGA 922
QY 814 CCAAAAAGCCATATAACAAGATTGTCTCACATTTGTGGCTGAACCGAGAAAGATCT 873
    |||||
Db 923 CAGCAGCCCAAGTGAATGCACTGTGTCTCATCTGTGTGTGAGCTGAGAAAGCTCT 982
QY 874 ATGCCATGCTGACCCCTACTGTCCCCGACAGTGAACATCAAGCCCTCACTACACTGTGTG 933
    |||||
Db 983 ATGCCATGCTGACCCCGCAGGCCCTGATGGGCACCTCCAGCCGTGGCTACCTCTGTG 1042
QY 934 ACTTGCCCGACCGAGATGGTGTATCATTTGATGGCGAAGCATATTCCAGGCTTCT 993
    |||||
Db 1043 ACCTCTTGACCGAGAGATTGTGTACCATCAAGCTGGGCCAAGACATCCAGGCTTCT 1102
QY 994 CCAGCTGTCCCTGGCGGACAGATGAGCCTTCTGCAGAGTCTTGATGGAATTTTGA 1053
    |||||
Db 1103 CATGCTGTGCTGTGTGACCAAGATGTCACTACTGCAGAGCGTGTGATGAGGTGCTGG 1162
QY 1054 TCCTGTGTGCTATACCGGTCTCTTCATTTGAGGATGAATGCTGTATGACAGCATT 1113
    |||||
Db 1163 TGCTGGGTGTGGCCAGCGCTCACTGCCACTGCAGAGATGAGCTTGCTTGAGGACT 1222
QY 1114 ATATATGACGAAGACAGTCCAATTAGCAGGCTTCTGATCTAAATAATGCTATCC 1173
    |||||
Db 1223 TAGTCTTGATGAAGAGGGGCAAGGCAAGTGCCTGGGGAACTGGGGCTGCCCTGC 1282
QY 1174 TGCAGCTGTTAAAGAAATACAGACATGAAGCTGGAAGAAAGAAATTTGTCACCTCA 1233
    |||||
Db 1283 TGCACCTAGTGGCGCGCTGCAGGCCCTCGCGCTGAGCGGAGAGATATGTTCTACTAA 1342
QY 1234 AAGCTATAGCTCTTGTCTAATTCAGACTCCATGCACATGAGAAT--GTTGAAGCGGTTC 1291
    |||||
Db 1343 AGGCTTGGCCCTTGCCAATTCAGACTCTGTGCACATGGAAGTGAAGCGCTGTGA 1402
QY 1292 GAAGCTTC-AGGATGTCTTACATGAAGCGCTGCAGATTATGAAGCTGGCCAG----- 1343
    |||||
Db 1403 GCAGCTGCAGAAGACTCCTGTGCAGAGCCCTGTGAGTATGAAGCGCGGCTGGCC 1462
QY 1344 -----CACATGGAAGACCCCTGTGAGCTGCAAGATGCTGATGACACTGCCACTCCTGA 1398
    |||||
Db 1463 CCGAGAGGGGTGCTGAGCGGCGGGCGGCGCAGGCTGTGCTACGCTACCGCTCCTCC 1522
QY 1399 GGCAGACCTCTACCAAGCGCGTGCAGCATTTCTACAACATCAAACTAGAAGCAAGTCC 1458
    |||||
Db 1523 GCCAGACAGCGGGCAAAAGTGTGCTGCCCATTTCTATGGGTGAAGCTGGAAGGCAAGTGC 1582
QY 1459 CAATGCACAAACTTTTGTGGAATGTTGAGGCCCAAGT 1498
Db 1583 CCATGCACAAAGCTGTTCTTGAGATGCTCGAGGCCATGAT 1622
```

```
RESULT 15
US-10-211-239-7
; Sequence 7, Application US/10211239
; Publication No. US20030103965A1
; GENERAL INFORMATION:
; APPLICANT: Jung, Birgit
; APPLICANT: Kraut, No. US20030103965Albert
; APPLICANT: Mueller, Stefan
; TITLE OF INVENTION: Method for Identifying Substances Which Positively
; TITLE OF INVENTION: Influence Inflammatory Conditions
; FILE REFERENCE: 0652.2340001
```

```

; CURRENT APPLICATION NUMBER: US/10/211,239
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/315,775
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: EP 01119003.0
; PRIOR FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-211-239-7

```

Query Match	13.5%	Score 401.6;	DB 15;	Length 2402;
Best Local Similarity	64.5%;	Pred. No. 7.4e-95;		
Matches 684;	Conservative	0;	Mismatches 334;	Indels 42;
				Gaps 4;

QY	481	TGCTCAACTCGATGCCCCAAGAGACTGTGTTT	TAGTGTGTGACATGCTTCTGGGTACC	540
Dp	563	TGCTCAGCTCCCTCGCCCAAGCGCTCTGCTT	GTGTGGGACGTGGCTCCGGCTACC	622
QY	541	ACTATGEEGGTAGCATCATGTGAAGCCTGCAAG	CATTCTTCAAGAGACAATTCAAGCA	600
Dp	623	ACTATGCTGTGGCATCTGTGAGGCTGCAAA	GCCTTCTTCAAGAGACCATCCAGGGGA	682
QY	601	ATATAGAATPACAGCTGCCCTGCCAAGATGA	TGTAATGCACAAGCCGACGCTAAT	660
Dp	683	GCAATGAGTACAGCTGTCCGGCTTCAACGAG	TGTGATATCAACAAGCGAGACGCAAG	742
QY	661	CCTGCCAGGCTTGCCGCTTCAATGAAGTGT	TAAAGTGGGCAATCTGAAAAGAGGGTGC	720
Dp	743	CCTGCCAGGCTTGCCGCTTCAACGAGTGC	CTCGGGTGGGCAATCTCAAGAGGAGTGC	802
QY	721	GTCTTGACAGAGTACGTGAGGTGCGCAGA	AGTACAAAGCGCATAGATGCGGAGACA	780
Dp	803	GCCTGGAACCGCGTCCGGGGTGGCGCGCA	AGATCAAGCGCGCGCGAGGTGACCCAC	862
QY	781	GCCCATACCTGAACCTCAGCTGCTGAGCC	-----AG	813
Dp	863	TGCCCTTCCGGGCCCCCTTCCCTGCTGGG	CCCCCTGCAATGCTGAGGCCCCCGAAGA	922
QY	814	CCAAAAAGCCATATTAACAAGATTGTCTCA	CATTGTGTGTGGCTGAACCGAGAGAATCT	873
Dp	923	CAGCAGCCCCAGTGAATGCACTGTGTCTCA	TCTGTCTGTGTGTGAGCCTGAGAGCTCT	982
QY	874	ATGCCATGCGCTGACCCCTACTGTCGCCG	ACAGTGAATCAAGCCCTCACTACACTGTGTG	933
Dp	983	ATGCCATGCGCTGACCCCGCAGGCCCTGAT	GGCACCTCCAGCCCGTGGCTACCTCTGTG	1042
QY	934	ACTTGGCCGACCGAGAGTGTGTGTTATCAT	TGATGGGCGAAGCATATTCAGGCTTCT	993
Dp	1043	ACCTCTTTGACCGAGAGATTGTGTCAACAT	CAGCTGGGCCAAGACATCCAGGCTTCT	1102
QY	994	CCACGCTGTCCCTGGCCGACCATGAGCCTT	CTGCAAGTCTTGATGGAATTTGA	1053
Dp	1103	CATCGCTGTGCTGTCTGACCAAGATGTCA	CTACTGCAAGCGTGTGATGAGGTGCTGG	1162
QY	1054	TCCTTGGTGTGATATACCGGCTCTTTCATT	TGAGGATGAACCTGTCTATGACAGCAT	1113
Dp	1163	TGCTGGGTGTGGCCAGCGCTCACTGCCACT	GACAGATGAGCTGGCCTTCTGTGAGACT	1222
QY	1114	ATATATATGACGAAGACCACTCCAAATTAG	CAGGCTTCTGATCTAAATATGCTATCC	1173
Dp	1223	TAGTCTGATGAAGAGGGGGACCGGACAGT	GTGGCTGGGGAACTGGGGCTGCC	1282
QY	1174	TGCAGCTGTAAAGAAATACAAGACATGA	AGCTGGAATAAAAGAAATTTGTCACCCTCA	1233
Dp	1283	TGCAACTAGTGCGGCGGCTGCAGGCCCTG	CGGCTGAGCGCAGAGAGATATGTTACTAA	1342
QY	1234	AAGCTATAGCTCTTGCTAATTCAGACTCCA	TGCACATAGAAGAT--GTTGAAGCCGTTCA	1291
Dp	1343	AGGCTTTGGCCCTTGCCAATTCAAGACTCT	GTGACATCGAAGATGAGCCGAGGCTGTGA	1402

OY		1292	GAACTTC-AGGATGTCTTACATGAAGCGCTGCAGATTATGAAGCTGGCCAG-----	1343
Db		1403	GCAGCTGCCAAGAAGCTCCTGCACGAGGCCCTGTGATATGAAGCCGGCCGGCTGGCC	1462
OY		1344	-----CACATGGAAGACCCTCGTGAGCTGGCAAGATGCTGATGACA CTGCCACTCTGA	1398
Db		1463	CCGAGGGG GTCTGAGCGCGCGGCGGGCAGGCTGCTCTCACGCTACCGCTCTCC	1522
OY		1399	GGCAGACCTCTAACCAAGCGCTGCAGCATTTCTAACATCAA ACTTAGAAGCAACTCC	1458
Db		1523	GCCA GACAGCGGGCA AAGTGTCTGGCCATTTCTATGGGATGAAGCTGGA GGCAAGTG C	1582
OY		1459	CAATGCACAAACTTTTTTTGGAAA TGTGGAGGCCAAAGT	1498
Db		1583	CCATGCACAAGCTGTCTTCTTGAGATGCTCGAGGCGCATGAT	1622

Search completed: August 10, 2004, 05:39:35  
Job time : 2420 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:35:05 ; Search time 54 Seconds  
(without alignments)  
2396.421 Million cell updates/sec

Title: US-10-054-841-4  
Perfect score: 2388  
Sequence: 1 MDSVELCLPESFSLHYEEL.....KLEGVPMHKLFLMLEAKV 458

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1:	geneseqp1980s:*
2:	geneseqp1990s:*
3:	geneseqp2000s:*
4:	geneseqp2001s:*
5:	geneseqp2002s:*
6:	geneseqp2003as:*
7:	geneseqp2003bs:*
8:	geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	2388	100.0	458	2	AAY03837 Human nuc
2	2388	100.0	458	3	AAB09965 Human ERR
3	2388	100.0	458	3	AAY83823 Human nuc
4	2388	100.0	458	7	ADC23486 Ligand bi
5	2388	100.0	458	7	ADD69757 Human ERR
6	2383	99.8	458	7	ADC23494 Mutant li
7	2382	99.7	458	7	ADC23493 Mutant li
8	2382	99.7	458	7	ADC23492 Mutant li
9	2374	99.4	458	7	ADC23495 Mutant li
10	2265	94.8	435	2	AAW61192 Steroid h
11	2265	94.8	435	3	AAY82263 Human oes
12	2235	93.6	436	7	ADD69755 Human ERR
13	2139	89.6	418	3	AAY83824 Human nuc
14	2122	88.9	418	2	AAY03838 Human nuc
15	2016.5	84.4	396	7	ADD69753 Human ERR
16	1803	75.5	433	3	AAB12970 Human oes
17	1803	75.5	508	2	AAY29467 Human SBP
18	1799	75.3	500	2	AAY03836 Human nuc
19	1799	75.3	500	3	AAY83822 Human nuc
20	1787	74.8	433	1	AAP80931 Sequence
21	1784	74.7	433	3	AAB09966 Human ERR
22	1259	52.7	486	5	ABB05347 Human nuc
23	1259	52.7	552	4	AAAM39411 Human pol
24	1258.5	52.7	519	6	ABP71695 Human est
25	1223.5	51.2	521	1	AAP80930 Sequence

26	1214.5	50.9	556	4	AAM41197 Human pol
27	1153.5	48.3	435	6	ABP97551 Amino aci
28	1146.5	48.0	240	6	ABR42765 Human oes
29	1073	44.9	516	5	ABP70149 Human NOV
30	938	39.3	376	5	AAE28940 Human nuc
31	739.5	31.0	484	4	ABB63251 Drosophil
32	739.5	31.0	484	6	AAE30102 Drosophil
33	739.5	31.0	519	6	AAE30101 Drosophil
34	726	30.4	587	5	ABB81600 Caiman cr
35	718	30.1	589	7	ADB99352 Fusion pr
36	715	29.9	581	5	ABB83791 Cnemidoph
37	713.5	29.9	595	5	AAU98991 Oestrogen
38	711.5	29.8	595	5	AAU98990 Oestrogen
39	707	29.6	595	4	AAG84512 Human oes
40	706	29.6	600	7	ADB99350 Fusion pr
41	704	29.5	595	2	AAY21626 Ligand bi
42	704	29.5	595	4	AAg84505 Human oes
43	704	29.5	595	4	AAg84513 Human oes
44	704	29.5	595	4	AAg84507 Human oes
45	704	29.5	595	5	ABB76378 Human nuc

ALIGNMENTS

RESULT 1  
AAY03837  
ID AAY03837 standard; protein; 458 AA.

AC AAY03837;  
AC 15-JUN-1999 (first entry)  
DT 15-JUN-1999 (first entry)  
DE Human nuclear receptor protein nNR2.  
DE Human nuclear receptor protein nNR2.  
DE Nuclear receptor; nNR1; nNR2; cell differentiation; human.  
KW Homo sapiens.  
XX Homo sapiens.  
XX WO9910367-A1.  
XX 04-MAR-1999.  
XX 27-AUG-1998; 98WO-US017826.  
XX 27-AUG-1997; 97US-0057090P.  
XX 21-OCT-1997; 97US-0062902P.  
XX 19-MAR-1998; 98US-0078633P.  
XX (MERI ) MERCK & CO INC.  
XX Chen F;  
XX WPI; 1999-190586/16.  
XX N-PSDB; AAX32266.  
XX Newly purified DNA polynucleotides encoding human nuclear trans-acting receptor proteins - useful in the diagnosis, treatment and prophylaxis of cell differentiation, development and physiological function.  
XX Claim 23; Page 46; 82pp; English.  
XX The invention relates to DNA molecules encoding human nuclear receptor (nNR) proteins nNR1 and nNR2. The nNR proteins (including mutants and/or fragments) form pharmaceutical compositions that are useful in the diagnosis, treatment and prophylaxis of cell differentiation, development and physiological function. The proteins are also useful for identifying downstream target genes and ligands regulating their activity. In particular, fusion constructs (especially glutathione S-transferase (GST) -nNR1 and/or GST-nNR2) expressing fusion proteins are useful in screening for (ant)agonists that are useful as modulators in cell differentiation, development and physiological function. Antibodies for the nNR proteins are useful for measuring the levels of these proteins. The present



CC sequence represents a human nNR2 protein  
XX  
SQ Sequence 458 AA;

Query Match 100.0%; Score 2388; DB 2; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9.2e-217;  
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEELLCRMNSNKRHIDSSCSFIFKTEPSSPASLTDVNHSPGG 60  
DB 1 MDSVELCLPESFSLHYEELLCRMNSNKRHIDSSCSFIFKTEPSSPASLTDVNHSPGG 60  
QY 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120  
DB 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120  
QY 121 NSMPKRLCLVCGDIAAGYHYGVAACEACKAFKKRTIQGNIEYSCPATNECEITKRRRKS 180  
DB 121 NSMPKRLCLVCGDIAAGYHYGVAACEACKAFKKRTIQGNIEYSCPATNECEITKRRRKS 180  
QY 181 QACRFMKCLKVGMKEGVRLDRVRGGRQKYKRRIDAENSPYLNPOLVQPAKKPYNKIVSH 240  
DB 181 QACRFMKCLKVGMKEGVRLDRVRGGRQKYKRRIDAENSPYLNPOLVQPAKKPYNKIVSH 240  
QY 241 LLVAEPEKIYAMPDPTVPDSDIKALITLCLDLADRELVIIGWAKHIPGSTLSLADQMSL 300  
DB 241 LLVAEPEKIYAMPDPTVPDSDIKALITLCLDLADRELVIIGWAKHIPGSTLSLADQMSL 300  
QY 301 LQSAWMEIILIGVVYRSLSFEDELVYADDYIMDEDQSKLAGLDLNNAILQLVKKYKSMK 360  
DB 301 LQSAWMEIILIGVVYRSLSFEDELVYADDYIMDEDQSKLAGLDLNNAILQLVKKYKSMK 360  
QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLQDVLEALQDYEAQOMEDPRRAGKMLMT 420  
DB 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLQDVLEALQDYEAQOMEDPRRAGKMLMT 420  
QY 421 LPLLROTSTKAVQHFYNIKLEGKVPMHKLFLEMLEAKV 458  
DB 421 LPLLROTSTKAVQHFYNIKLEGKVPMHKLFLEMLEAKV 458

RESULT 2  
AAB09965 standard; protein; 458 AA.

XX AAB09965;  
AC  
XX  
DT 19-OCT-2000 (first entry)  
XX  
DE Human ERGamma protein.  
XX  
KW Human; ERGamma; brain; estrogen-related receptor gamma; cytostatic;  
KW lipid metabolism; fatty acid synthesis; antiarteriosclerotic; treatment;  
KW drug development; diabetes.  
XX  
OS Homo sapiens.

XX  
PN WO200026365-A1.  
XX  
PD 11-MAY-2000.

XX  
PF 02-NOV-1999; 99WO-JP006097.  
XX  
PR 04-NOV-1998; 98JP-00313194.  
XX  
PA (KAZU-) KAZUSA DNA RES INST FOUND.  
XX (TAIS ) TAISHO PHARM CO LTD.  
XX  
PI Ohara O, Nagase T, Nomura N, Takayama K, Toyoda H, Yoshimoto M;

XX  
DR WPI; 2000-365614/31.  
XX N-PSDB; AAA40078, AAA40079.  
XX

PT An estrogen-related receptor gamma protein with lipid metabolism  
PT regulatory and fatty acid synthesis functions, and its encoding gene,  
PT useful as drugs and in developing drugs for treatment of e.g.  
PT arteriosclerosis and diabetes.

PS Claim 1a; Page 22-24; 38pp; Japanese.

CC This invention describes a novel estrogen-related receptor gamma  
CC (ERGamma) protein (I) and its variants maintaining receptor activity,  
CC having lipid metabolism regulatory and fatty acid synthesis functions.  
CC The product of the invention has cytostatic and antiarteriosclerotic  
CC activity. The protein and gene are useful as drugs and in developing  
CC drugs for treatment of e.g. arteriosclerosis and diabetes. The protein is  
CC obtained by cloning human brain origin-originated cDNA library. This sequence  
CC represents the human brain ERGamma protein which is described in the  
CC method of the invention

XX  
SQ Sequence 458 AA;

Query Match 100.0%; Score 2388; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9.2e-217;  
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEELLCRMNSNKRHIDSSCSFIFKTEPSSPASLTDVNHSPGG 60  
DB 1 MDSVELCLPESFSLHYEELLCRMNSNKRHIDSSCSFIFKTEPSSPASLTDVNHSPGG 60  
QY 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120  
DB 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120  
QY 121 NSMPKRLCLVCGDIAAGYHYGVAACEACKAFKKRTIQGNIEYSCPATNECEITKRRRKS 180  
DB 121 NSMPKRLCLVCGDIAAGYHYGVAACEACKAFKKRTIQGNIEYSCPATNECEITKRRRKS 180  
QY 181 QACRFMKCLKVGMKEGVRLDRVRGGRQKYKRRIDAENSPYLNPOLVQPAKKPYNKIVSH 240  
DB 181 QACRFMKCLKVGMKEGVRLDRVRGGRQKYKRRIDAENSPYLNPOLVQPAKKPYNKIVSH 240  
QY 241 LLVAEPEKIYAMPDPTVPDSDIKALITLCLDLADRELVIIGWAKHIPGSTLSLADQMSL 300  
DB 241 LLVAEPEKIYAMPDPTVPDSDIKALITLCLDLADRELVIIGWAKHIPGSTLSLADQMSL 300  
QY 301 LQSAWMEIILIGVVYRSLSFEDELVYADDYIMDEDQSKLAGLDLNNAILQLVKKYKSMK 360  
DB 301 LQSAWMEIILIGVVYRSLSFEDELVYADDYIMDEDQSKLAGLDLNNAILQLVKKYKSMK 360  
QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLQDVLEALQDYEAQOMEDPRRAGKMLMT 420  
DB 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLQDVLEALQDYEAQOMEDPRRAGKMLMT 420  
QY 421 LPLLROTSTKAVQHFYNIKLEGKVPMHKLFLEMLEAKV 458  
DB 421 LPLLROTSTKAVQHFYNIKLEGKVPMHKLFLEMLEAKV 458

RESULT 3  
AAY83823

ID AAY83823 standard; protein; 458 AA.

XX AAY83823;  
AC

XX  
DT 05-JUL-2000 (first entry)  
XX

XX Human nuclear receptor nNR2.

XX Human nuclear receptor protein-2; nNR2; physiological function;  
KW cell development and differentiation controller; gene expression.  
XX  
OS Homo sapiens.

XX  
XX US6054295-A.  
XX

PD 25-APR-2000.  
XX  
PF 26-AUG-1998; 98US-00141000.  
XX  
PR 27-AUG-1997; 97US-0057090P.  
PR 21-OCT-1997; 97US-0062922P.  
PR 19-MAR-1998; 98US-0078633P.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Chen F;  
XX  
DR WPI; 2000-328352/28.  
DR N-PSDB; AAA09801.  
XX  
XX  
PT New polynucleotide encoding human nuclear receptor protein-1 is useful  
PT for screening compounds that acts as modulators of cell differentiation,  
PT cell development and physiological function.  
XX  
XX  
PS Disclosure; Fig 5; 58pp; English.  
XX  
XX This sequence represents the human nuclear receptor protein-2 (nNR2). The  
CC protein is a cell development and differentiation controller which is  
CC useful for screening compounds that act as modulators of cell  
CC differentiation, cell development and physiological functions or for  
CC treating and diagnosing disorders associated with altered expression of  
CC the DNA  
XX  
SQ Sequence 458 AA;  
  
Query Match 100.0%; Score 2388; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9.2e-217;  
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MDSVELCLPESFSLHYEELLCRMENKDRHIDSSCSFFIKTEPSSPASLTDSVNHSPGG 60  
Db 1 MDSVELCLPESFSLHYEELLCRMENKDRHIDSSCSFFIKTEPSSPASLTDSVNHSPGG 60  
QY 61 SSDASGSYSTMNGHONGLDSPPLYPAPILGSGPVRKLYDDCSSTIVEDPQTKCEYML 120  
Db 61 SSDASGSYSTMNGHONGLDSPPLYPAPILGSGPVRKLYDDCSSTIVEDPQTKCEYML 120  
QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACAKAFKRTIQGNIIEYSCPATNECEITKRRKSC 180  
Db 121 NSMPKRLCLVCGDIASGYHYGVAACEACAKAFKRTIQGNIIEYSCPATNECEITKRRKSC 180  
QY 181 QACRFMKCLKVGMKEGVRIDRVRGGRQYKRRIDAENSPYLNQLVQPAKKPYNKIVSH 240  
Db 181 QACRFMKCLKVGMKEGVRIDRVRGGRQYKRRIDAENSPYLNQLVQPAKKPYNKIVSH 240  
QY 241 LLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVIIGWAKHIPGFSTLSLADQMSL 300  
Db 241 LLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVIIGWAKHIPGFSTLSLADQMSL 300  
QY 301 LOSAWMEILLIGVYVYRSLSEFDELVYADYIMDEDQSKLAGLLDNNAILQLVKKYSMK 360  
Db 301 LOSAWMEILLIGVYVYRSLSEFDELVYADYIMDEDQSKLAGLLDNNAILQLVKKYSMK 360  
QY 361 LEKEEFVTLKAIALANDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420  
Db 361 LEKEEFVTLKAIALANDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420  
QY 421 LPLLRQTSITAVQHFYNYIKLEGKVPMEKLFLEMLEAKV 458  
Db 421 LPLLRQTSITAVQHFYNYIKLEGKVPMEKLFLEMLEAKV 458  
  
RESULT 4  
ADC23486  
ID ADC23486 standard; protein; 458 AA.  
XX  
AC ADC23486;  
XX

DT 18-DEC-2003 (First entry)  
XX  
DE Ligand binding domain of the oestrogen related receptor 3 (ERR3) protein.  
XX  
XX oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;  
KW ligand binding domain; fertility; birth control; bone remodelling;  
KW cancer; protein coordinate data.  
XX  
OS Unidentified.  
XX  
XX WO2003064468-A2.  
PN  
PD 07-AUG-2003.  
XX  
XX 30-JAN-2003; 2003WO-EP000959.  
PF  
XX 31-JAN-2002; 2002US-0352551P.  
PR  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX Moras D, Renaud J, Greschik H, Wurtz J;  
PI  
XX WPI; 2003-663467/62.  
DR  
DR N-PSDB; ADC23488.  
XX  
XX New peptide fragment, useful for screening compounds that are agonists or  
PT antagonists of the transcriptional-activating activity of the estrogen-  
PT related receptor 3 (ERR3).  
XX  
XX  
PS Claim 1; SEQ ID NO 1; 259pp; English.  
XX  
XX This invention relates to novel peptide fragments that have an agonistic  
CC or antagonistic effect on the transcriptional-activating activity of the  
CC oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen  
CC related receptor gamma protein. Specifically, the peptide fragment of the  
CC invention comprises a ligand binding domain of ERR3, which becomes  
CC functionally active when fused to a protein containing a DNA binding  
CC domain and mimics the transcriptional-activating activity of the complete  
CC ERR3 protein. As such, the peptide fragment can be used in a screening  
CC method or to design and select compounds that affect ERR3 activity.  
CC Furthermore, the agonists and antagonists of ERR3 are biologically active  
CC compounds that can be used to modulate the oestrogenic response on  
CC fertility, birth control, bone remodelling, breast and prostate cancer.  
CC This polypeptide sequence is the ligand binding pocket of the ERR3  
CC protein of the invention.  
XX  
SQ Sequence 458 AA;  
  
Query Match 100.0%; Score 2388; DB 7; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9.2e-217;  
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MDSVELCLPESFSLHYEELLCRMENKDRHIDSSCSFFIKTEPSSPASLTDSVNHSPGG 60  
Db 1 MDSVELCLPESFSLHYEELLCRMENKDRHIDSSCSFFIKTEPSSPASLTDSVNHSPGG 60  
QY 61 SSDASGSYSTMNGHONGLDSPPLYPAPILGSGPVRKLYDDCSSTIVEDPQTKCEYML 120  
Db 61 SSDASGSYSTMNGHONGLDSPPLYPAPILGSGPVRKLYDDCSSTIVEDPQTKCEYML 120  
QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACAKAFKRTIQGNIIEYSCPATNECEITKRRKSC 180  
Db 121 NSMPKRLCLVCGDIASGYHYGVAACEACAKAFKRTIQGNIIEYSCPATNECEITKRRKSC 180  
QY 181 QACRFMKCLKVGMKEGVRIDRVRGGRQYKRRIDAENSPYLNQLVQPAKKPYNKIVSH 240  
Db 181 QACRFMKCLKVGMKEGVRIDRVRGGRQYKRRIDAENSPYLNQLVQPAKKPYNKIVSH 240  
QY 241 LLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVIIGWAKHIPGFSTLSLADQMSL 300  
Db 241 LLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVIIGWAKHIPGFSTLSLADQMSL 300  
QY 301 LOSAWMEILLIGVYVYRSLSEFDELVYADYIMDEDQSKLAGLLDNNAILQLVKKYSMK 360

Db 301 LQSAWMEILILGVVYRSLSFEDELVYADYIMDEDQSKLAGLIDLINNAIILQLVKKYKSMK 360

QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420

Db 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420

QY 421 LPILRQSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

Db 421 LPILRQSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

RESULT 5

ADD69757

ID ADD69757 standard; protein; 458 AA.

XX AC ADD69757;

XX DT 15-JAN-2004 (first entry)

XX DE Human ERR gamma 3-related protein - SEQ ID 6.

XX KW nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;

XX KW oestrogen receptor; ER; thyroid hormone; TR; human.

XX OS Homo sapiens.

XX PN WO2003080831-A1.

XX PD 02-OCT-2003.

XX PF 25-MAR-2003; 2003WO-JP003611.

XX PR 25-MAR-2002; 2002JP-00084560.

XX PA (FUJII ) FUJISAWA PHARM CO LTD.

XX PI Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;

XX DR WPI; 2003-779262/73.

XX DR N-PSDB; ADD69756.

XX PT Polynucleotides encoding nuclear receptors, and the encoded proteins, useful as diagnostic agents, and for identification of agents that affect receptor activity.

XX PS Claim 9; SEQ ID NO 6; 148pp; Japanese.

XX CC The invention relates to novel nuclear receptor ERR (oestrogen receptor-related receptor) gamma 3 polynucleotides. The polynucleotides of the invention may be useful for diagnosis of disorders caused by abnormal nuclear receptor activity, particularly those related to abnormal oestrogen receptor (ER), ERR or thyroid hormone receptor (TR) activity. Furthermore, the polynucleotides and proteins may be useful for evaluating agents that affect the activity of nuclear receptors. The current sequence is that of the human ERR gamma 3-related protein of the invention.

XX SQ Sequence 458 AA;

Query Match 100.0%; Score 2388; DB 7; Length 458;

Best Local Similarity 100.0%; Pred. No. 9.2e-217;

Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEEELICRMSNKRHRIDSSCSFFIKTEPSSPASLTDVNMHSPGG 60

Db 1 MDSVELCLPESFSLHYEEELICRMSNKRHRIDSSCSFFIKTEPSSPASLTDVNMHSPGG 60

QY 61 SSDASGSYSTMNGHONGLDSPPLYPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120

Db 61 SSDASGSYSTMNGHONGLDSPPLYPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120

QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFKFKRTIQGNIIEYSCPATNECETTKRRKSC 180

Db 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFKFKRTIQGNIIEYSCPATNECETTKRRKSC 180

QY 181 QACRFMKCLKVGMLEKGVRLDRVRGGRQKRRIDAENSPLYNPOLVQPAKPPYNKIVSH 240

Db 181 QACRFMKCLKVGMLEKGVRLDRVRGGRQKRRIDAENSPLYNPOLVQPAKPPYNKIVSH 240

QY 241 LIVAEPKITYAMPDPTVPDSDIKALITLTLCDLADRELVTIIGWAKHIFGSTLSLADQMSL 300

Db 241 LIVAEPKITYAMPDPTVPDSDIKALITLTLCDLADRELVTIIGWAKHIFGSTLSLADQMSL 300

QY 301 LQSAWMEILILGVVYRSLSFEDELVYADYIMDEDQSKLAGLIDLINNAIILQLVKKYKSMK 360

Db 301 LQSAWMEILILGVVYRSLSFEDELVYADYIMDEDQSKLAGLIDLINNAIILQLVKKYKSMK 360

QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420

Db 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420

QY 421 LPILRQSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

Db 421 LPILRQSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

RESULT 6

ADC23494

ID ADC23494 standard; protein; 458 AA.

XX AC ADC23494;

XX DT 18-DEC-2003 (first entry)

XX DE Mutant ligand binding domain of the ERR3 protein - A272L.

XX KW oestrogen related receptor 3; ERR3; oestrogen related receptor gamma; ligand binding domain; fertility; birth control; bone remodelling; cancer; protein coordinate data; mutant; mutetin.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Misc-difference 272

FT /note= "wild type Ala substituted for Leu"

XX PN WO2003064468-A2.

XX PD 07-AUG-2003.

XX PF 30-JAN-2003; 2003WO-EP000959.

XX PR 31-JAN-2002; 2002US-0352551P.

XX PA (CNRS ) CNRS CENT NAT RECH SCT.

XX PI Moras D, Renaud J, Greschik H, Wurtz J;

XX DR WPI; 2003-663467/62.

XX PT New peptide fragment, useful for screening compounds that are agonists or antagonists of the transcriptional-activating activity of the estrogen-related receptor 3 (ERR3).

XX PS Claim 49; Page; 259pp; English.

XX CC This invention relates to novel peptide fragments that have an agonistic or antagonistic effect on the transcriptional-activating activity of the oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen related receptor gamma protein. Specifically, the peptide fragment of the invention comprises a ligand binding domain of ERR3, which becomes functionally active when fused to a protein containing a DNA binding domain and mimics the transcriptional-activating activity of the complete ERR3 protein. As such, the peptide fragment can be used in a screening method or to design and select compounds that affect ERR3 activity.



CC Furthermore, the agonists and antagonists of ERR3 are biologically active  
CC compounds that can be used to modulate the oestrogenic response on  
CC fertility, birth control, bone remodeling, breast and prostate cancer.  
CC This polypeptide sequence is a mutant protein (A272L) comprising the  
CC ligand binding pocket of the ERR3 protein of the invention. NOTE: This  
CC sequence is not given in the specification but is derived from  
CC information given in the claim 49.

XX  
SQ Sequence 458 AA;

Query Match 99.8%; Score 2383; DB 7; Length 458;  
Best Local Similarity 99.8%; Pred. No. 2.7e-216;  
Matches 457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEELLCRMSNKRHIDSSCSSFIKTEPSSPASLTDSVNHSPGG 60  
DB 1 MDSVELCLPESFSLHYEELLCRMSNKRHIDSSCSSFIKTEPSSPASLTDSVNHSPGG 60  
QY 61 SSDASGSYSTMNGHNGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYML 120  
DB 61 SSDASGSYSTMNGHNGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYML 120  
QY 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIEYSCPATNECEITKRRKSC 180  
DB 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIEYSCPATNECEITKRRKSC 180  
QY 181 QACRFMKCLKVGMKEGVRLDRVRGGRQKYKRRIDAENSPYLPQLVQPAKKPYNKIVSH 240  
DB 181 QACRFMKCLKVGMKEGVRLDRVRGGRQKYKRRIDAENSPYLPQLVQPAKKPYNKIVSH 240  
QY 241 LLVAEPEKIYAMPDPYVDSIDIKALITLCLDLADRELVIIGWAKHIPGFSTLSLADQMSL 300  
DB 241 LLVAEPEKIYAMPDPYVDSIDIKALITLCLDLADRELVIIGWAKHIPGFSTLSLADQMSL 300  
QY 301 IQSAWMEILILGVVYRSLSFEDELVYADDYIMDEQSKLAGLLDNNAILQLVKYYKSMK 360  
DB 301 IQSAWMEILILGVVYRSLSFEDELVYADDYIMDEQSKLAGLLDNNAILQLVKYYKSMK 360  
QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420  
DB 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420  
QY 421 LPLLRQSTSTKAVQHFYNIKLEGKVPMHKLFLEMLEAKV 458  
DB 421 LPLLRQSTSTKAVQHFYNIKLEGKVPMHKLFLEMLEAKV 458

RESULT 7  
ADC23493  
ID ADC23493 standard; protein; 458 AA.  
XX  
AC ADC23493;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Mutant ligand binding domain of the ERR3 protein - A272F.  
XX  
KM oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;  
KM ligand binding domain; fertility; birth control; bone remodeling;  
KW cancer; protein coordinate data; mutant; mutein.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 272 /note= "Wild type Ala substituted for Phe"  
XX  
PN WO2003064468-A2.  
XX  
XX 07-AUG-2003.  
XX  
PD 30-JAN-2003; 2003WO-EP000959.  
PF  
XX

PR 31-JAN-2002; 2002US-0352551P.  
XX  
XX  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX  
PI Moras D, Renaud J, Greschik H, Wurtz J;  
XX  
XX  
DR WPI; 2003-663467/62.

PT New peptide fragment, useful for screening compounds that are agonists or  
PT antagonists of the transcriptional-activating activity of the estrogen-  
PT related receptor 3 (ERR3).  
XX  
PS Claim 49; Page; 259pp; English.

XX  
CC This invention relates to novel peptide fragments that have an agonistic  
CC or antagonistic effect on the transcriptional-activating activity of the  
CC oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen  
CC related receptor gamma protein. Specifically, the peptide fragment of the  
CC invention comprises a ligand binding domain of ERR3, which becomes  
CC functionally active when fused to a protein containing a DNA binding  
CC domain and mimics the transcriptional-activating activity of the complete  
CC ERR3 protein. As such, the peptide fragment can be used in a screening  
CC method or to design and select compounds that affect ERR3 activity.  
CC Furthermore, the agonists and antagonists of ERR3 are biologically active  
CC compounds that can be used to modulate the oestrogenic response on  
CC fertility, birth control, bone remodeling, breast and prostate cancer.  
CC This polypeptide sequence is a mutant protein (A272F) comprising the  
CC ligand binding pocket of the ERR3 protein of the invention. NOTE: This  
CC sequence is not given in the specification but is derived from  
CC information given in the claim 49.

XX  
SQ Sequence 458 AA;

Query Match 99.7%; Score 2382; DB 7; Length 458;  
Best Local Similarity 99.8%; Pred. No. 3.4e-216;  
Matches 457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEELLCRMSNKRHIDSSCSSFIKTEPSSPASLTDSVNHSPGG 60  
DB 1 MDSVELCLPESFSLHYEELLCRMSNKRHIDSSCSSFIKTEPSSPASLTDSVNHSPGG 60  
QY 61 SSDASGSYSTMNGHNGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYML 120  
DB 61 SSDASGSYSTMNGHNGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYML 120  
QY 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIEYSCPATNECEITKRRKSC 180  
DB 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIEYSCPATNECEITKRRKSC 180  
QY 181 QACRFMKCLKVGMKEGVRLDRVRGGRQKYKRRIDAENSPYLPQLVQPAKKPYNKIVSH 240  
DB 181 QACRFMKCLKVGMKEGVRLDRVRGGRQKYKRRIDAENSPYLPQLVQPAKKPYNKIVSH 240  
QY 241 LLVAEPEKIYAMPDPYVDSIDIKALITLCLDLADRELVIIGWAKHIPGFSTLSLADQMSL 300  
DB 241 LLVAEPEKIYAMPDPYVDSIDIKALITLCLDLADRELVIIGWAKHIPGFSTLSLADQMSL 300  
QY 301 IQSAWMEILILGVVYRSLSFEDELVYADDYIMDEQSKLAGLLDNNAILQLVKYYKSMK 360  
DB 301 IQSAWMEILILGVVYRSLSFEDELVYADDYIMDEQSKLAGLLDNNAILQLVKYYKSMK 360  
QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420  
DB 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKIQDVLHEALQDYEAGQHMEDPRRAGKMLMT 420  
QY 421 LPLLRQSTSTKAVQHFYNIKLEGKVPMHKLFLEMLEAKV 458  
DB 421 LPLLRQSTSTKAVQHFYNIKLEGKVPMHKLFLEMLEAKV 458

RESULT 8  
ADC23492  
ID ADC23492 standard; protein; 458 AA.

XX AC ADC23492;  
XX DT 18-DEC-2003 (first entry)  
XX DE Mutant ligand binding domain of the ERR3 protein - F435L.  
XX KW oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;  
KW ligand binding domain; fertility; birth control; bone remodelling;  
KW cancer; protein coordinate data; mutant; mutein.  
XX OS Unidentified.  
XX FH Key Location/Qualifiers  
FT Misc-difference 435 /note= "wild type Phe substituted for leu"  
XX PN WO2003064468-A2.  
XX PD 07-AUG-2003.  
XX PE 30-JAN-2003; 2003WO-EP000959.  
XX PR 31-JAN-2002; 2002US-0352551P.  
XX PA (CNRS ) CNRS CENT NAT RECH SCT.  
XX PI Moras D, Renaud J, Greschik H, Wurtz J;  
XX DR WPI; 2003-663467/62.  
XX PT New peptide fragment, useful for screening compounds that are agonists or  
PT antagonists of the transcriptional-activating activity of the estrogen-  
PT related receptor 3 (ERR3).  
XX PS Claim 49; Page; 259pp; English.  
XX CC This invention relates to novel peptide fragments that have an agonistic  
CC or antagonistic effect on the transcriptional-activating activity of the  
CC oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen  
CC related receptor gamma protein. Specifically, the peptide fragment of the  
CC invention comprises a ligand binding domain of ERR3, which becomes  
CC functionally active when fused to a protein containing a DNA binding  
CC domain and mimics the transcriptional-activating activity of the complete  
CC ERR3 protein. As such, the peptide fragment can be used in a screening  
CC method or to design and select compounds that affect ERR3 activity.  
CC Furthermore, the agonists and antagonists of ERR3 are biologically active  
CC compounds that can be used to modulate the oestrogenic response on  
CC fertility, birth control, bone remodelling, breast and prostate cancer.  
CC This polypeptide sequence is a mutant protein (F435L) comprising the  
CC ligand binding pocket of the ERR3 protein of the invention. NOTE: This  
CC sequence is not given in the specification but is derived from  
CC information given in the claim 49.  
XX CC  
XX CC  
SQ Sequence 458 AA;  
  
Query Match 99.7%; Score 2382; DB 7; Length 458;  
Best Local Similarity 99.8%; Pred. No. 3.4e-216;  
Matches 457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEELLCRMNKNDRHIDSSCSSFIKTEPSSPASLTDSVNAHSPGG 60  
Db 1 MDSVELCLPESFSLHYEELLCRMNKNDRHIDSSCSSFIKTEPSSPASLTDSVNAHSPGG 60  
  
QY 61 SSDASGSYSTMNGHONGLDSPPLYPSAPILGSGSPVRKLYDDCSSTIVEDPQTCKEYML 120  
Db 61 SSDASGSYSTMNGHONGLDSPPLYPSAPILGSGSPVRKLYDDCSSTIVEDPQTCKEYML 120  
  
QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFKRTIQGNIEYSCPATNECETIKRRKSC 180  
Db 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFKRTIQGNIEYSCPATNECETIKRRKSC 180  
  
QY 181 QACRFMKCLKVGMKEGVRLDRVRGGRÖKYKRRIDAENSPYLNQVLQPAKKPYNKIVSH 240

Db 181 QACRFMKCLKVGMKEGVRLDRVRGGRÖKYKRRIDAENSPYLNQVLQPAKKPYNKIVSH 240  
QY 241 LLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTYIGWAKHIPGFSTLSLADQMSL 300  
Db 241 LLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTYIGWAKHIPGFSTLSLADQMSL 300  
QY 301 LQSAWEILILGVYRSLSEDELVYADDYIMDEDOQSLAGLLDLNNAIQLVKKYYKSMK 360  
Db 301 LQSAWEILILGVYRSLSEDELVYADDYIMDEDOQSLAGLLDLNNAIQLVKKYYKSMK 360  
QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVÖKLQDVLHEALQDYEAQÖHMEDPRRAGKMLMT 420  
Db 361 LEKEEFVTLKAIALANSDSMHIEDVEAVÖKLQDVLHEALQDYEAQÖHMEDPRRAGKMLMT 420  
QY 421 LPLLRÖSTKAVÖHFYNIKLEGVPMHKLFLLEMLEAKV 458  
Db 421 LPLLRÖSTKAVÖHLVNIKLEGVPMHKLFLLEMLEAKV 458  
  
RESULT 9  
ID ADC23495 standard; protein; 458 AA.  
XX AC ADC23495;  
XX DT 18-DEC-2003 (first entry)  
XX DE Mutant ligand binding domain of the ERR3 protein - L345I/ F435L/ F450L.  
XX KW oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;  
KW ligand binding domain; fertility; birth control; bone remodelling;  
KW cancer; protein coordinate data; mutant; mutein.  
XX OS Unidentified.  
XX FH Key Location/Qualifiers  
FT Misc-difference 345 /note= "wild type leu substituted for ile"  
FT Misc-difference 435 /note= "wild type Phe substituted for leu"  
FT Misc-difference 450 /note= "wild type Phe substituted for leu"  
XX PN WO2003064468-A2.  
XX PD 07-AUG-2003.  
XX PE 30-JAN-2003; 2003WO-EP000959.  
XX PR 31-JAN-2002; 2002US-0352551P.  
XX PA (CNRS ) CNRS CENT NAT RECH SCT.  
XX PI Moras D, Renaud J, Greschik H, Wurtz J;  
XX DR WPI; 2003-663467/62.  
XX PT New peptide fragment, useful for screening compounds that are agonists or  
PT antagonists of the transcriptional-activating activity of the estrogen-  
PT related receptor 3 (ERR3).  
XX PS Claim 49; Page; 259pp; English.  
XX CC This invention relates to novel peptide fragments that have an agonistic  
CC or antagonistic effect on the transcriptional-activating activity of the  
CC oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen  
CC related receptor gamma protein. Specifically, the peptide fragment of the  
CC invention comprises a ligand binding domain of ERR3, which becomes  
CC functionally active when fused to a protein containing a DNA binding  
CC domain and mimics the transcriptional-activating activity of the complete  
CC ERR3 protein. As such, the peptide fragment can be used in a screening  
CC method or to design and select compounds that affect ERR3 activity.

CC Furthermore, the agonists and antagonists of ERR3 are biologically active  
CC compounds that can be used to modulate the oestrogenic response on  
CC fertility, birth control, bone remodeling, breast and prostate cancer.  
CC This polypeptide sequence is a mutant protein (U345I/ F435L/ F450L)  
CC comprising the ligand binding pocket of the ERR3 protein of the  
CC invention. NOTE: This sequence is not given in the specification but is  
CC derived from information given in the claim 49.

XX Sequence 458 AA;

Query Match 99.4%; Score 2374; DB 7; Length 458;  
Best Local Similarity 99.3%; Pred. No. 1.9e-215;  
Matches 455; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEEELLCRMNKNDRHIDSSCSFITEPSSPASLTDVNHSPGG 60  
Db 1 MDSVELCLPESFSLHYEEELLCRMNKNDRHIDSSCSFITEPSSPASLTDVNHSPGG 60  
QY 61 SSDASGYSSTMGHONGLDSPPLYPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYML 120  
Db 61 SSDASGYSSTMGHONGLDSPPLYPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYML 120  
QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFKKRTIQGNIEYSCPATNECEITKRRKSC 180  
Db 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFKKRTIQGNIEYSCPATNECEITKRRKSC 180  
QY 181 QACRFMKCLKVGMKEGVRLDRVGGGRQYKRRIDAENSPYLNQVLVQPAKKPYNKIVSH 240  
Db 181 QACRFMKCLKVGMKEGVRLDRVGGGRQYKRRIDAENSPYLNQVLVQPAKKPYNKIVSH 240  
QY 241 LIVAEEPKIYAMPDPTVPDSIDIKALITLCLDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300  
Db 241 LIVAEEPKIYAMPDPTVPDSIDIKALITLCLDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300  
QY 301 LOSAMEILILGVVYRSLSFEDLVYADDYIMDEDQSKLAGLLDNNAILQLVKKYKSMK 360  
Db 301 LOSAMEILILGVVYRSLSFEDLVYADDYIMDEDQSKLAGLLDNNAILQLVKKYKSMK 360  
QY 361 LEKEEFVTLKAIALANSDSMTHIEDEVAVOKLQDVLHEALQDYEAQOMEDPRRAGKMLMT 420  
Db 361 LEKEEFVTLKAIALANSDSMTHIEDEVAVOKLQDVLHEALQDYEAQOMEDPRRAGKMLMT 420  
QY 421 LPLLRQSTSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458  
Db 421 LPLLRQSTSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

RESULT 10  
AAW61192  
ID AAW61192 standard; protein; 435 AA.

XX AAW61192;  
XX 07-DEC-1998 (first entry)  
XX Steroid hormone receptor homologue HE8AB36.  
DE HE8AB36; steroid hormone receptor; human; transcription factor;  
KW inflammation; arthritis; autoimmune disease; diabetes;  
KW transplant rejection; graft versus host disease; cancer;  
KW reproductive disorder; obesity; atherosclerosis; gyrate atrophy; therapy;  
KW diagnosis.  
XX Homo sapiens.  
OS EP866127-A2.  
PN 23-SEP-1998.  
PD 13-MAR-1998;  
PF 13-MAR-1998;  
XX 17-MAR-1997;  
PR 16-FEB-1998;  
AC 98GB-00003289.

XX (SMIK ) SMITHKLINE BEECHAM PLC.  
XX Mathias SL;  
XX WPI: 1998-482963/42.  
DR N-PSDB; AAV47645.

PT New isolated polypeptide(s) are steroid hormone receptor homologues -  
PT used for treating inflammation, arthritis, auto-immune disease, diabetes,  
PT transplant rejection etc.

PS Claim 4; Page 12; 20pp; English.

XX This polypeptide comprises human HE8AB36, a member of the nuclear hormone  
CC receptor family of polypeptides that are ligand regulated transcription  
CC factors involved in the regulation of cellular homeostasis and  
CC differentiation via the modulation of gene transcription. Its amino acid  
CC sequence was deduced from an isolated cDNA clone (see AAV47645). HE8AN36  
CC polypeptides and polynucleotides may be used in the treatment of chronic  
CC and acute inflammation, arthritis, autoimmune diseases, diabetes,  
CC cancer, obesity, atherosclerosis, gyrate atrophy and other visual  
CC disorders. The invention also relates to methods for identifying agonists  
CC and antagonists/inhibitors, and for treating conditions associated with  
CC HE8AN36 imbalance using such compounds. Diagnostic assays for detecting  
CC diseases associated with inappropriate HE8AN36 activity or levels are  
CC also provided

SQ Sequence 435 AA;

Query Match 94.8%; Score 2265; DB 2; Length 435;  
Best Local Similarity 100.0%; Pred. No. 3.7e-205;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MSNKNDRHIDSSCSFITEPSSPASLTDVNHSPGGSSDASGYSSTMGHONGLDSP 83  
Db 1 MSNKNDRHIDSSCSFITEPSSPASLTDVNHSPGGSSDASGYSSTMGHONGLDSP 83  
QY 84 LYPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYMLNSMPKRLCLVCGDIASGYHYGVA 143  
Db 61 LYPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYMLNSMPKRLCLVCGDIASGYHYGVA 120  
QY 144 SCEACKAFKKRTIQGNIEYSCPATNECEITKRRRKSQACRFMKCLKVGMKEGVRLDRV 203  
Db 121 SCEACKAFKKRTIQGNIEYSCPATNECEITKRRRKSQACRFMKCLKVGMKEGVRLDRV 180  
QY 204 RGGROKYKRRIDAENSPYLNQVLVQPAKKPYNKIVSHLIVAEEPKIYAMPDPTVPDSID 263  
Db 181 RGGROKYKRRIDAENSPYLNQVLVQPAKKPYNKIVSHLIVAEEPKIYAMPDPTVPDSID 240  
QY 264 ALTLCLDLADRELVTIIGWAKHIPGFSTLSLADQMSLLOSAMEILILGVVYRSLSFED 323  
Db 241 ALTLCLDLADRELVTIIGWAKHIPGFSTLSLADQMSLLOSAMEILILGVVYRSLSFED 300  
QY 324 LUYADDYIMDEDQSKLAGLLDNNAILQLVKKYKSMKLEKEEFVTLKAIALANSDSMTH 383  
Db 301 LUYADDYIMDEDQSKLAGLLDNNAILQLVKKYKSMKLEKEEFVTLKAIALANSDSMTH 360  
QY 384 DVEAVQKLQDVLHEALQDYEAQOMEDPRRAGKMLMTLPLLRQSTSTKAVQHFYNIKLE 443  
Db 361 DVEAVQKLQDVLHEALQDYEAQOMEDPRRAGKMLMTLPLLRQSTSTKAVQHFYNIKLE 420  
QY 444 VPMHKLFLMLEAKV 458  
Db 421 VPMHKLFLMLEAKV 435

RESULT 11  
AAV82263  
ID AAV82263 standard; protein; 435 AA.  
XX AAV82263;  
AC AAV82263;



XX 14-JUN-2000 (first entry)  
XX  
XX Human oestrogen related receptor 3 SEQ ID NO:1.  
DE  
XX  
XX Human; ligand-combining; oestrogen related receptor; ERR3; diagnosis;  
KW inflammation; cancer; osteoporosis; diabetes; renal disease.  
XX  
OS Homo sapiens.  
XX  
XX JP2000041681-A.  
PN  
XX 15-FEB-2000.  
PD  
XX 31-JUL-1998; 98JP-00217933.  
PF  
XX 31-JUL-1998; 98JP-00217933.  
XX  
PR 31-JUL-1998; 98JP-00217933.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
PA  
XX  
XX WPI; 2000-295782/26.  
DR  
XX N-PSDB; AAZ95741.  
DR  
A new protein comprising a 435 amino acid sequence.  
PT  
XX  
XX Claim 1; Page 18-19; 22pp; Japanese.  
PS  
XX  
XX The present sequence represents a human oestrogen related receptor 3  
CC (ERR3) protein which has ligand-combining activity. The polynucleotide  
CC sequence encoding ERR3, and fragments of it, can be used for the  
CC diagnosis and the treatment of inflammations, cancers, osteoporosis,  
CC diabetes or renal diseases  
XX  
XX  
SQ Sequence 435 AA;

Query Match 94.8%; Score 2265; DB 3; Length 435;  
Best Local Similarity 100.0%; Pred. No. 3.7e-205;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 24 MSNKRHIDSSCSFIKTEPSPASLTDVNHSPGSSSDASGSYSTMNGHQLDSP 83  
Db 1 MSNKRHIDSSCSFIKTEPSPASLTDVNHSPGSSSDASGSYSTMNGHQLDSP 60  
QY 84 LYPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 143  
Db 61 LYPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 120  
QY 144 SCEACKAFKKRTIQGNI EYSCPATNECEITKRRRKSQACRFMKCLKVGMKEGVRLDRV 203  
Db 121 SCEACKAFKKRTIQGNI EYSCPATNECEITKRRRKSQACRFMKCLKVGMKEGVRLDRV 180  
QY 204 RGRQKYKRRIDAENSPYLNQLVQPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK 263  
Db 181 RGRQKYKRRIDAENSPYLNQLVQPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK 240  
QY 264 ALTTLCDLADRELTVIIGWAKHIFGFSTLSLADQMSILQSAWMEIILIGVYRSLSFDE 323  
Db 241 ALTTLCDLADRELTVIIGWAKHIFGFSTLSLADQMSILQSAWMEIILIGVYRSLSFDE 300  
QY 324 LVYADDYIMDEDQSKLAGLLDNNAILQLVKYKSMKLEKEEFVTLKAIALANSDSMHE 383  
Db 301 LVYADDYIMDEDQSKLAGLLDNNAILQLVKYKSMKLEKEEFVTLKAIALANSDSMHE 360  
QY 384 DVEAVQKLQDVLHEALQDYEAGQHEDPRRAGKMLMTPLIRQSTTKAVQHFYNIKLEGK 443  
Db 361 DVEAVQKLQDVLHEALQDYEAGQHEDPRRAGKMLMTPLIRQSTTKAVQHFYNIKLEGK 420  
QY 444 VPMHKLFLMLEAKV 458  
Db 421 VPMHKLFLMLEAKV 435

RESULT 12

ADD69755  
ID ADD69755 standard; protein; 436 AA.  
XX  
XX AC ADD69755;  
XX  
XX DT 15-JAN-2004 (first entry)  
XX  
XX DE Human ERR gamma 3-related protein - SEQ ID 4.  
XX  
XX KW nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;  
KW oestrogen receptor; ER; thyroid hormone; TR; human.  
XX  
OS Homo sapiens.  
XX  
XX PN WO2003080831-A1.  
XX  
XX PD 02-OCT-2003.  
XX  
XX PF 25-MAR-2003; 2003WO-JP003611.  
XX  
XX PR 25-MAR-2002; 2002JP-00084560.  
XX  
XX (FUJI ) FUJISAWA PHARM CO LTD.  
XX  
XX PA Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;  
XX  
XX PI WPI; 2003-779262/73.  
XX  
XX DR N-PSDB; ADD69754.  
XX  
XX PT Polynucleotides encoding nuclear receptors, and the encoded proteins,  
PT useful as diagnostic agents, and for identification of agents that affect  
PT receptor activity.  
XX  
XX  
PS Claim 9; SEQ ID NO 4; 148pp; Japanese.  
XX

XX  
XX The invention relates to novel nuclear receptor ERR (oestrogen receptor-  
CC related receptor) gamma 3 polynucleotides. The polynucleotides of the  
CC invention may be useful for diagnosis of disorders caused by abnormal  
CC nuclear receptor activity, particularly those related to abnormal  
CC oestrogen receptor (ER), ERR or thyroid hormone receptor (TR) activity.  
CC Furthermore, the polynucleotides and proteins may be useful for  
CC evaluating agents that affect the activity of nuclear receptors. The  
CC current sequence is that of the human ERR gamma 3-related protein of the  
CC invention.  
XX  
XX  
SQ Sequence 436 AA;

Query Match 93.6%; Score 2235; DB 7; Length 436;  
Best Local Similarity 98.9%; Pred. No. 2.6e-202;  
Matches 430; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 24 MSNKRHIDSSCSFIKTEPSPASLTDVNHSPGSSSDASGSYSTMNGHQLDSP 83  
Db 1 MSNKRHIDSSCSFIKTEPSPASLTDVNHSPGSSSDASGSYSTMNGHQLDSP 60  
QY 84 LYPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 143  
Db 61 LYPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 120  
QY 144 SCEACKAFKKRTIQGNI EYSCPATNECEITKRRRKSQACRFMKCLKVGMKEGVRLDRV 203  
Db 121 SCEACKAFKKRTIQGNI EYSCPATNECEITKRRRKSQACRFMKCLKVGMKEGVRLDRV 180  
QY 204 RGRQKYKRRIDAENSPYLNQLVQPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK 263  
Db 181 RGRQKYKRRIDAENSPYLNQLVQPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK 240  
QY 264 ALTTLCDLADRELTVIIGWAKHIFGFSTLSLADQMSILQSAWMEIILIGVYRSLSFDE 323  
Db 241 ALTTLCDADRELTVIIGWAKHIFGFSTLSLADQMSILQSAWMEIILIGVYRSLSFDE 300  
QY 324 LVYADDYIMDEDQSKLAGLLDNNAILQLVKYKSMKLEKEEFVTLKAIALANSDSMHE 383  
Db 301 LVYADDYIMDEDQSKLAGLLDNNAILQLVKYKSMKLEKEEFVTLKAIALANSDSMHE 360

Db 301 LVVADDYIMDEDQSKLAGLLDNNAILQLVKKYSMKLEKEEFVTLKALALANSDSMHIE 360

QY 384 DVEAVOKLQDVLHEALQDYEAQGHMEDPRRAGKMLTLLRQSTSTKAVQHFYNTKLEGK 443

Db 361 DVEAVOKLQDVLHEALQDYEAQGHMEDPRRAGKMLTLLRQSTSTKAVQHFYNTKLEGK 420

QY 444 VEMMKLFLEMLEAKV 458

Db 421 VEMMKLFLEMLEAKV 435

RESULT 13

AAy83824

ID AAY83824 standard; protein; 418 AA.

XX AC AAY83824;

XX DT 05-JUL-2000 (first entry)

XX DE Human nuclear receptor nNR2-1 variant protein.

XX KM Human nuclear receptor protein 2-1; nNR2-1; physiological function; cell development and differentiation controller; gene expression.

XX OS Homo sapiens.

XX PN US6054295-A.

XX PD 25-APR-2000.

XX PF 26-AUG-1998; 98US-00141000.

XX PR 27-AUG-1997; 97US-0057090P.

XX PR 21-OCT-1997; 97US-0062922P.

XX PR 19-MAR-1998; 98US-0078633P.

XX PA (MERI ) MERCK & CO INC.

XX PI Chen F;

XX DR WPI; 2000-328352/28.

XX DR N-PSDB; AAA09802.

XX PT New polynucleotide encoding human nuclear receptor protein-1 is useful for screening compounds that acts as modulators of cell differentiation, cell development and physiological function.

XX PS Disclosure; Fig 8; 58pp; English.

XX CC This sequence represents the variant human nuclear receptor protein 2-1 (nNR2-1). The coding gene differs from the nNR2 gene (AAA09801) by the insertion of 2 nucleotides at position 1352. The insertion results in the shifting of the reading frame and the creation of a termination codon 33 nucleotides downstream of the insertion site. The translated protein has a C-terminal truncation. The nNR proteins are cell development and differentiation controllers which are useful for screening compounds that act as modulators of cell differentiation, cell development and physiological functions or for treating and diagnosing disorders associated with altered expression of the DNA

XX SQ Sequence 418 AA;

Query Match 89.6%; Score 2139; DB 3; Length 418;

Best Local Similarity 100.0%; Pred. No. 3e-193;

Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEEELCRMENKDRHIDSSCSFFIKTEPSSPASLTDSVNHSPGG 60

Db 1 MDSVELCLPESFSLHYEEELCRMENKDRHIDSSCSFFIKTEPSSPASLTDSVNHSPGG 60

QY 61 SSDAGSYSTMNGHQGLDSPPLVPSAPILGGSPVRKLYDDCSSTIVEDPQTKCEYML 120

Db 61 SSDAGSYSTMNGHQGLDSPPLVPSAPILGGSPVRKLYDDCSSTIVEDPQTKCEYML 120

QY 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFERKTIQGNIEYSCPATNECETKRRKSC 180

Db 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFERKTIQGNIEYSCPATNECETKRRKSC 180

QY 181 QACRFMKCLKVGMLEKGVRLDRVRGGRQKKYRRIDAENSPYLNPOLVQPAKKPYNKIVSH 240

Db 181 QACRFMKCLKVGMLEKGVRLDRVRGGRQKKYRRIDAENSPYLNPOLVQPAKKPYNKIVSH 240

QY 241 LLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300

Db 241 LLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300

QY 301 LOSAWMEILILGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDNNAILQLVKKYSMK 360

Db 301 LOSAWMEILILGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDNNAILQLVKKYSMK 360

QY 361 LEKEEFVTLKALALANSDSMHIEDVEAVOKLQDVLHEALQDYEAQGHME 409

Db 361 LEKEEFVTLKALALANSDSMHIEDVEAVOKLQDVLHEALQDYEAQGHME 409

RESULT 14

AAy03838

ID AAY03838 standard; protein; 418 AA.

XX AC AAY03838;

XX DT 15-JUN-1999 (first entry)

XX DE Human nuclear receptor protein nNR2-1.

XX KM Nuclear receptor; nNR1; nNR2; cell differentiation; human.

XX OS Homo sapiens.

XX PN WO910367-A1.

XX PD 04-MAR-1999.

XX PF 27-AUG-1998; 98WO-US017826.

XX PR 27-AUG-1997; 97US-0057090P.

XX PR 21-OCT-1997; 97US-0062902P.

XX PR 19-MAR-1998; 98US-0078633P.

XX PA (MERI ) MERCK & CO INC.

XX PI Chen F;

XX DR WPI; 1999-190586/16.

XX DR N-PSDB; AAX32267.

XX PT Newly purified DNA polynucleotides encoding human nuclear trans-acting receptor proteins - useful in the diagnosis, treatment and prophylaxis of cell differentiation, development and physiological function.

XX PS Disclosure; Fig 8; 82pp; English.

XX CC The invention relates to DNA molecules encoding human nuclear receptor (nNR) proteins nNR1 and nNR2. The nNR proteins (including mutants and/or fragments) form pharmaceutical compositions that are useful in the diagnosis, treatment and prophylaxis of cell differentiation, development and physiological function. The proteins are also useful for identifying downstream target genes and ligands regulating their activity. In particular, fusion constructs (especially glutathione S-transferase (GST) -nNR1 and/or GST-nNR2) expressing fusion proteins are useful in screening for (ant)agonists that are useful as modulators in cell differentiation, development and physiological function. Antibodies for the nNR proteins are useful for measuring the levels of these proteins. The present sequence represents a human nNR2-1 protein

XX SQ Sequence 418 AA;

Query Match 88.9%; Score 2122; DB 2; Length 418;  
Best Local Similarity 99.5%; Pred. No. 1.2e-191;  
Matches 407; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEEELCRMNSKDRHIDSSCSSFIKTEPSSPASLTDSVNHSPGG 60  
Db 1 MDSVELCLPESFSLHYEEELCRMNSKDRHIDSSCSSFIKTEPSSPASLTDSVNHSPGG 60

QY 61 SSDASGSYSTMNGHONGLDSPPLPSAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120  
Db 61 SSDASGSYSTMNGHONGLDSPPLPSAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120

QY 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIIEYSCPATNECEITKRRKSC 180  
Db 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIIEYSCPATNECEITKRRKSC 180

QY 181 QACRFMKCLKVGMKEGVRLLDRVGRGQKYKRRIDAENSPYLNQVLQPAKKPYNKIVSH 240  
Db 181 QACRFMKCLKVGMKEGVRLLDRVGRGQKYKRRIDAENSPYLNQVLQPAKKPYNKIVSH 240

QY 241 LTVABEPEKIYAMPDPTVPDSIDIKALLTLLCDLADRELVIIGWAKHIFGSTLSLADQMSL 300  
Db 241 LTVABEPEKIYAMPDPTVPDSIDIKALLTLLCDLADRELVIIGWAKHIFGSTLSLADQMSL 300

QY 301 LQSAWMEILLGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDNNAILQLVKYKYSMK 360  
Db 301 LQSAWMEILLGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDNNAILQLVKYKYSMK 360

QY 361 LEKEEFVTLKAILALANSDSMHIEDVEAVQKLQDVLHEALQDYEAGQHME 409  
Db 361 LEKEEFVTLKAILALANSDSMHIEDVEAVQKLQDVLHEALQDYEAGQHME 409

RESULT 15  
ADD69753  
ID ADD69753 standard; protein; 396 AA.

XX ADD69753;  
AC  
XX 15-JAN-2004 (first entry)  
DT  
XX  
DE Human ERR gamma 3-related protein - SEQ ID 2.  
XX  
KW nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;  
KW oestrogen receptor; ER; thyroid hormone; TR; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2003080831-A1.  
XX  
XX 02-OCT-2003.  
PD  
XX 25-MAR-2003; 2003WO-JP003611.  
PF  
XX 25-MAR-2002; 2002JP-00084560.  
PR  
XX (FUJII ) FUJISAWA PHARM CO LTD.  
PA  
XX Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;  
PI  
XX WPI; 2003-779262/73.  
DR N-PSDB; ADD69752.  
XX  
XX  
PT Polynucleotides encoding nuclear receptors, and the encoded proteins,  
PT useful as diagnostic agents, and for identification of agents that affect  
PT receptor activity.  
XX  
XX Claim 1; SEQ ID NO 2; 148pp; Japanese.  
PS  
XX  
CC The invention relates to novel nuclear receptor ERR (oestrogen receptor-  
CC related receptor) gamma 3 polynucleotides. The polynucleotides of the  
CC invention may be useful for diagnosis of disorders caused by abnormal

CC nuclear receptor activity, particularly those related to abnormal  
CC oestrogen receptor (ER), ERR or thyroid hormone receptor (TR) activity.  
CC Furthermore, the polynucleotides and proteins may be useful for  
CC evaluating agents that affect the activity of nuclear receptors. The  
CC current sequence is that of the human ERR gamma 3-related protein of the  
CC invention.  
XX  
SQ Sequence 396 AA;

Query Match 84.4%; Score 2016.5; DB 7; Length 396;  
Best Local Similarity 90.8%; Pred. No. 1.1e-181;  
Matches 395; Conservative 0; Mismatches 1; Indels 39; Gaps 1;

QY 24 MSNKRHIDSSCSSFIKTEPSSPASLTDSVNHSPGSSSDASGSYSTMNGHONGLDSP 83  
Db 1 MSNKRHIDSSCSSFIKTEPSSPASLTDSVNHSPGSSSDASGSYSTMNGHONGLDSP 60

QY 84 LYPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 143  
Db 61 LYPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 120

QY 144 SCEACKAFFKRTIQGNIIEYSCPATNECEITKRRKSCQACRFMKCLKVGMKEGVRLLDRV 203  
Db 121 SCEACKAFFKRTIQ-----GVRLLDRV 141

QY 204 RGGQKYKRRIDAENSPYLNQVLQPAKKPYNKIVSHLLVABEPEKIYAMPDPTVPDSIDIK 263  
Db 142 RGGQKYKRRIDAENSPYLNQVLQPAKKPYNKIVSHLLVABEPEKIYAMPDPTVPDSIDIK 201

QY 264 ALLTLCDLADRELVIIGWAKHIFGSTLSLADQMSLLQSAWMEILLGVVYRSLSFED 323  
Db 202 ALLTLCDLADRELVIIGWAKHIFGSTLSLADQMSLLQSAWMEILLGVVYRSLSFED 261

QY 324 LVYADDYIMDEDQSKLAGLLDNNAILQLVKYKYSMKLEKEEFVTLKAILALANSDSMHIE 383  
Db 262 LVYADDYIMDEDQSKLAGLLDNNAILQLVKYKYSMKLEKEEFVTLKAILALANSDSMHIE 321

QY 384 DVEAVQKLQDVLHEALQDYEAGQHMEBPRAGKMLMTLLPLRQSTKAVQHFYNIKLEEK 443  
Db 322 DVEAVQKLQDVLHEALQDYEAGQHMEBPRAGKMLMTLLPLRQSTKAVQHFYNIKLEEK 381

QY 444 VPMHKLFLFMLBAKV 458  
Db 382 VPMHKLFLFMLBAKV 396

Search completed: August 3, 2004, 12:40:31  
Job time : 55 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:38:36 ; Search time 17 Seconds  
(without alignments)  
2591.513 Million cell updates/sec

Title: US-10-054-841-4  
Perfect score: 2388  
Sequence: 1 MDSVELCLPESFSLHYEEL.....KLEGVPMHKLFLLEMLAKV 458

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1784	74.7	433	2 B29345	steroid hormone re
2	1748	73.2	433	2 S58087	estrogen receptor
3	1223.5	51.2	521	2 A29345	steroid hormone re
4	718	30.1	589	1 QRCHE	estrogen receptor
5	716.5	30.0	600	1 QQRTE	estrogen receptor
6	713.5	29.9	595	2 I47140	estradiol receptor
7	710	29.7	599	1 QRMSE	estrogen receptor
8	702	29.4	595	1 QRHUE	estrogen receptor
9	694	29.1	586	1 QRXLB	estrogen receptor
10	674	28.2	535	2 S58224	oestrogen receptor
11	674	28.2	620	2 T10423	estrogen receptor
12	666.5	27.9	574	2 A37197	estrogen receptor
13	662	27.7	701	2 S64737	80K estrogen recep
14	648	27.1	530	2 JCS939	estrogen receptor
15	644.5	27.0	477	2 S71400	estrogen receptor
16	634	26.5	503	2 JW0046	estrogen receptor
17	623.5	26.1	441	2 I50515	retinoid X recepto
18	609.5	25.5	470	2 D41977	retinoid receptor
19	600	25.1	488	2 C41977	retinoid receptor
20	598	25.0	467	2 A43781	retinoid-X-recepto
21	596.5	25.0	467	2 S26668	retinoid acid rece
22	591.5	24.8	467	2 A47278	retinoid X recepto
23	589	24.7	462	1 S09592	retinoid X recepto
24	577.5	24.2	463	2 B41727	retinoid-X recepto
25	575	24.1	379	2 I50514	retinoid X recepto
26	573	24.0	533	2 S37781	retinoid X recepto
27	572.5	24.0	463	2 S2670	retinoid acid rece
28	561	23.5	451	2 A41651	retinoid acid rece
29	560	23.5	446	2 A34418	H-2 region II bind

30	560	23.5	448	2 D41727	retinoid X recepto
31	560	23.5	452	2 S47633	RXR protein - Afri
32	560	23.5	520	2 I84718	RXR-beta1 isoform
33	553.5	23.2	410	2 S26669	retinoid acid rece
34	551	23.1	422	2 I50516	retinoid X recepto
35	550.5	23.1	438	2 I50517	retinoid X recepto
36	545	22.8	476	2 B41977	retinoid acid rece
37	528.5	22.1	458	2 A34714	retinoid acid rece
38	527	22.1	454	1 A33903	retinoid acid rece
39	525	22.0	560	2 S27874	steroid hormone re
40	524	21.9	454	2 S06124	retinoid acid rece
41	522	21.9	442	2 A38592	retinoid acid rece
42	515.5	21.6	403	2 S35334	steroid receptor p
43	514	21.5	543	2 A32693	steroid receptor p
44	511	21.4	458	2 S06123	retinoid acid rece
45	509	21.3	443	1 C35991	retinoid acid rece

ALIGNMENTS

RESULT 1  
B29345  
steroid hormone receptor ERR2 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 20-Sep-1999  
C:Accession: B29345  
R:Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.  
Nature 331, 91-94, 1988  
A:Title: Identification of a new class of steroid hormone receptors.  
A:Reference number: A29345; MUID:88122546; PMID:3267207  
A:Accession: B29345  
A:Molecule type: mRNA  
A:Residues: 1-433 <GIG>  
A:Cross-references: GB:X51417; NID:g36610; PIDN:CAA35779.1; PID:g36611; GB:Y00290  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: DNA binding; steroid hormone receptor; transcription regulation; zinc finger  
F:101-352/Domain: erba transforming protein homology <ERBA>  
F:103-123/Region: zinc finger  
F:139-163/Region: zinc finger

Query Match	Query	Score	Length	DB	Length
Best Local Similarity	74.7%;	Score 1784;	DB 2;	Length 433;	
Matches 333;	Conservative 51;	Mismatches 49;	Indels 2;	Gaps 1;	
QY	24 MSNKRHDIDSSCSSFIKTEPSSPASITDSVNHSPGSSSDASGSYSTMNGHONGDSDSP	83			
DB	1 MSSEDRHIGSSCGSFIKTEPSSPSGTDALSHSPSSGSDASGFGWALGTHANGDSDSP	60			
QY	84 LYPSPATLGGSGPVRKLYDDCSSSTIVEDPQTKCEYMLNSMPKRLCTVCGDIASGYHYGVA	143			
DB	61 MPAGAGL--GGNPCRKSYEDCTSGIMEDSAIKCEYMLNAIPKRLCLVCGDIASGYHYGVA	118			
QY	144 SCEACKAFPRKRTIQGNIYSCPATNECETITRRRKSQACRFMKCLKVGMKEGVRIDRV	203			
DB	119 SCEACKAFPRKRTIQGNIYSCPATNECETITRRRKSQACRFMKCLKVGMKEGVRIDRV	178			
QY	204 RGGROKYKRRIIDAEENSPYLNQOLVQPAKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK	263			
DB	179 RGGROKYKRRIIDAEENSPYLNQOLVQPAKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK	238			
QY	264 ALTTLCIDLADRELVIIGWAKHIFGFSSTLADQMSLQSAWMEILILGVYRSLFEDE	323			
DB	239 ALTTLCIDLADRELVIIGWAKHIFGFSSTLADQMSLQSAWMEILILGVYRSLFEDE	298			
QY	324 LVYADDYIMDEQSLAGLLDNNAILQVKKYKSMKLEKEFVTLKALANSDSMHIE	383			
DB	299 LAYAEYIMDEHSLVGLLELYRAILQVRRYKKLKVEKEFEVMTKALANSDSMYIE	358			
QY	384 DVEAVQKLQVLHEALQDYEAQGHMEDPRRAGKMLMTPLPLRQSTKAVQHFYNIKLEGK	443			
DB	359 NLEAVQKLQVLHEALQDYELSQRHEEPRRAGKMLMTPLPLRQSTKAVQHFYSVKLQGX	418			

QY 444 VPMHKLFLMLEAKV 458  
Db 419 VPMHKLFLMLEAKV 433

RESULT 2

estrogen receptor related protein - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
C/Accession: S58087  
R/Pettersson, K.; Svensson, K.; Mattsson, R.; Carlsson, B.; Ohlsson, R.; Berkenstam, A.  
submitted to the EMBL Data Library, July 1995  
A/Description: Expression of a novel member of estrogen response element-binding nuclear  
A/Reference number: S58087  
A/Accession: S58087  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-433 <PET>  
A/Cross-references: EMBL:X89594; NID:g914855; PIDN:CAA61755.1; PID:g914856  
C/Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C/Keywords: steroid hormone receptor; zinc finger  
F;101-352/Domain: erba transforming protein homology <ERBA>

Query Match 73.2%; Score 1748; DB 2; Length 433;  
Best Local Similarity 74.9%; Pred. No. 2.5e-121;  
Matches 326; Conservative 54; Mismatches 53; Indels 2; Gaps 1;

QY 24 MSNKDRHIDSSCSFIKTEPSSPASLTDVNMHSPGSSSDASGSYSTMNGHQGLDSP 83  
Db 1 MSSEDRHLCSSCGSFIKTEPSSPSGGLDALSHSPSGSSDASGFGIALSTHANGLDSP 60  
QY 84 LYPSPALIGSGSPVRKLYDDCSSITYEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143  
Db 61 MFAGAGL--CGNPCRKSYEDCTSGIMEDSAIKCEYMLNAIPKRLCLVCGDIASGYHYGVA 118  
QY 144 SCEACKAFPKRTIQGNIEYSCPAINECEITKRRKSCQACREMKCLKVGMLKEGVRLD 203  
Db 119 SCEACKAFPKRTIQGNIEYNCPAINECEITKRRKSCQACREMKCLKVGMLKEGVRLD 178  
QY 204 RGGROKKYRRIDAENSPYLPQVOPAKKPYNKIVSHLVAEPEKIYAMPDPTVPDS 263  
Db 179 RGGROKKYRRIDSENSPYLPISPAKKPLTKRIVSNLGVQDKLYAMPNDIPEGDIK 238  
QY 264 ALTTLCDLADRELVTIGWAKHIPGFSSTLSADQMSLQSAWMEILLGVYRSLSEFEDE 323  
Db 239 ALTTLCDLADRELVTFLINWAKHIPGFSSTLTGQMSLQSAWMEILLGVYRSLSEFEDE 298  
QY 324 LVVADYIMDEDQSKLAGLLDNNAILQLVKYKSMKLEKEEFTLLKAIALANSDSM 383  
Db 299 LAYAEYIMDEHSRLVGLLDYRALQLVRYKYLKVEKEEFMLIKALALANSDSMYIE 358  
QY 384 DVEAVQKLDVILHEALQDYEAAGQHMEDPRRAGKMLMTPLRLQTSKAVQHFYNIKLE 443  
Db 359 NLEAVQKLDVILHEALQDYELSQRHHEPRRAGKMLMTPLRLQTSKAVQHFYSVKLG 418  
QY 444 VPMHKLFLMLEAKV 458  
Db 419 VPMHKLFLMLEAKV 433

RESULT 3

A29345  
steroid hormone receptor ERRI precursor - human  
N/Alternate names: estrogen-related receptor  
C/Species: Homo sapiens (man)  
C/Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 20-Sep-1999  
C/Accession: A29345; A49074  
R/Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.  
Nature 331, 91-94, 1988  
A/Title: Identification of a new class of steroid hormone receptors.  
A/Reference number: A29345; MUID:88122546; PMID:3267207  
A/Accession: A29345

A/Molecule type: mRNA  
A/Residues: 1-521 <GIG>  
A/Cross-references: EMBL:X51416; NID:g36608; PIDN:CAA35778.1; PID:g36609; EMBL:Y00290  
R/Wiley, S.R.; Kraus, R.J.; Zuo, F.; Murray, E.E.; Loritz, K.; Mertz, J.E.  
Genes Dev. 7, 2206-2219, 1993  
A/Title: SV40 early-to-late switch involves titration of cellular transcriptional repres;  
A/Reference number: A49074; MUID:94040741; PMID:8224847  
A/Accession: A49074

A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 166-169, 'X', 171-173 <WIL>  
C/Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C/Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; zi  
F;174-434/Domain: erba transforming protein homology <ERBA>  
F;176-196/Region: zinc finger  
F;212-236/Region: zinc finger

Query Match 51.2%; Score 1223.5; DB 2; Length 521;  
Best Local Similarity 54.0%; Pred. No. 1.7e-82;  
Matches 252; Conservative 66; Mismatches 84; Indels 65; Gaps 9;

QY 22 CRMSNKDRHIDSSCS-----FIKTEPSSPASLTDVNMHSPGSSD-----A 64  
Db 85 CPTALPEQVTSAMSSQVVGIEPLTYIKAPASP-----DSFGSSETEPEPVALA 135  
QY 65 SG-SYSTMNGHQGLDSPPLYPSPALIGSGSPVRKLYDDCSSITYEDPQTKCEYMLNSM 123  
Db 136 PGPAFTCLPGHKEEDGE-----GAGP-----GEQGGKLVLSL 171  
QY 124 PKRLCLVCGDIASGYHYGVASCEACKAFPKRTIQGNIEYSCPAINECEITKRRKSCQAC 183  
Db 172 PKRLCLVCGDVASGYHYGVASCEACKAFPKRTIQGSIEYSCPAINECEITKRRKACQAC 231  
QY 184 RFMKCLKVGMLKEGVRLDVRVGGROKKYRRIDAENSPYLPQVOP-----AKKPY 234  
Db 232 RFTKCLRVGMLKEGVRLDVRVGGROKKYRRPEVDPLFPFGPFPAGPLAVAAGPRKTAAPV 291  
QY 235 NKIVSHLVAEPEKIYAMPDPTVPDSIKALTTLCDLADRELVTIGWAKHIPGFSSTLS 294  
Db 292 NALVSHLIVBEKLYAMPDPAGDGLPAVATLCDLFDREIVVTISWAKSIPGSSLSL 351  
QY 295 ADQMSLQSAWMEILLGVYRSLSEFEDELVYADYIMDEDQSKLAGLLDNNAILQLVK 354  
Db 352 SDQMSVLIQSVMEVLVLGVAQRPDLQDELAFADLVLDDEGARAAGLGELGALLQLVR 411  
QY 355 KYKSMKLEKEEFTLLKAIALANSDSMHIEDVEAV-QKLQDVILHEALQDYEAQ---HME 409  
Db 412 RLQALRLEREYVLLKALALANSDSVHIEDPRLWSCEKLLHEALLEYEAAGRAGPGGA 471  
QY 410 DPRRAGKMLMTPLRLQTSKAVQHFYNIKLEGVPMHKLFLMLEA 456  
Db 472 ERRRRGRLLTLPLRLQTAGKVLAHFYGVKLEGVPMHKLFLMLEA 518

RESULT 4

ORCHE  
estrogen receptor - chicken  
C/Species: Gallus gallus (chicken)  
C/Date: 06-Mar-1992 #sequence\_revision 14-Jul-1994 #text\_change 22-Jun-1999  
C/Accession: A40914; S07192  
R;Maxwell, B.L.; McDonnell, D.P.; Conneely, O.M.; Schulz, T.Z.; Greene, G.L.; O'Malley, I  
Mol. Endocrinol. 1, 25-35, 1987  
A/Title: Structural organization and regulation of the chicken estrogen receptor.  
A/Reference number: A40914; MUID:88318621; PMID:2901032  
A/Accession: A40914  
A/Status: not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-589 <MAX>  
R;Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.; Bornert, J.M.; Chambon, P.  
EMBO J. 5, 891-897, 1986  
A/Title: The chicken oestrogen receptor sequence: homology with v-erbA and the human oest  
A/Reference number: S07192; MUID:86247578; PMID:3755102  
A/Accession: S07192

A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-256,'E',258-589 <KRU>  
A:Cross-references: EMBL:X03805; NID:g63378; PIDN:CAA27433.1; PID:g63380  
C:Comment: The steroid hormones and their receptors are involved in the regulation of eu  
C:Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly  
complex appears to recognize discrete DNA sequences upstream of transcriptional start sit  
C:Superfamily: estrogen receptor; erba transforming protein homology  
C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone recep  
F:1-173/Domain: amino-terminal <NH2>  
F:174-265/Domain: DNA binding #status predicted <DNA>  
F:177-450/Domain: erba transforming protein homology <ERBA>  
F:177-200/Region: zinc finger CCCC motif  
F:213-235/Region: zinc finger CCCC motif  
F:250-265/Region: nuclear location signal  
F:294-546/Domain: steroid binding #status predicted <STB>  
F:179,182,196,199/Binding site: zinc (Cys) #status predicted  
F:215,221,231,234/Binding site: zinc (Cys) #status predicted  
F:230,299/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 30.1%; Score 718; DB 1; Length 589;  
Best Local Similarity 36.8%; Pred. No. 4.3e-45;  
Matches 170; Conservative 90; Mismatches 134; Indels 68; Gaps 14;

QY 38 FIKTEPSSPASTDSVNHS---PGSSDASGSYSTMNGHNGLDSPPLY--PSAPITG 92  
Db 104 FLQTAP-----QLSPFIHHSSQVPPYLENEGFSF-----GMREA--APPAFYRPPS----- 148  
QY 93 GSGPYVKLYDDCCSTIVEDPQTKE---YMLNSMPKRLCLVCGDIASGYHYGVASCEAC 148  
Db 149 -----DNRRHSIRERMSSITNEKSLSWESTKETRYCAVCNDYASGYHYGVWSCEGC 199  
QY 149 KAFEKRTIQNIEYSCPATNECEITKRRKSCQACREFMKCLKYGMKEGVRLLDRVRGGR- 207  
Db 200 KAFEKRTIQGHNDYMCPATNQCTIDKRRKSCQACRLKCYEYGVMMKGIRKDR-RGGRM 258  
QY 208 QKXKRRIIDAENSP-----YINPOLVQPAKK-----PYNKIVSHLLVAEP 246  
Db 259 MKQKRQREQDSRNGEASSTELRAPLTWTSPLVVKNKNKNSPALSLTAEQWVSALLAEAP 318  
QY 247 EKIYAMPDPTVPDSDIKALTTLCDLADRELVIIGWAKHIPGFSTLSLADQMSLLQSAWM 306  
Db 319 PIVYSEYDNRPFENASMTTLTNLADRELVMINWAKRVPGFVDTLLHDQVHLLCEAWL 378  
QY 307 EILILGVYRSLSEFEDLVYADDYIMEDQSK-LAGLLDLNNAITQLVKKYKSMKLEKEE 365  
Db 379 EILMIGLVWSMEHPGKLLFAPNLLDRNQSKVEGWEIFDMLLATARFRMNLQGE 438  
QY 366 FVTLKATAIAN-----SDSMHIEDVEAVQKLQDVLEALQDYEAGQHM--EDPRRA 414  
Db 439 FVCLKSIILNSGVYTFLSSTLKSLEERYIHRVLDKITDTLLHMAKSGLSLQQQHRRL 498  
QY 415 GKMLMTPLIRQSTSTKAVOHFYNIKLEGVPMHKLFLLEMLEA 456  
Db 499 AQLLLILSHIRMSNKGMEHLYNMCKKNVPLLYDLLLEMLDA 540

RESULT 5  
ORSTE  
estrogen receptor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Sep-1991 #sequence\_revision 14-Jul-1994 #text\_change 22-Jun-1999  
C:Accession: S07379; S16731  
R:Koike, S.; Sakai, M.; Muramatsu, M.  
Nucleic Acids Res. 15, 2499-2513, 1987  
A:Title: Molecular cloning and characterization of rat estrogen receptor cDNA.  
A:Reference number: S07379; MUID:87174780; PMID:3031601  
A:Accession: S07379  
A:Molecule type: mRNA  
A:Residues: 1-600 <KOI>  
A:Cross-references: EMBL:Y00102; NID:g56110; PIDN:CAA68287.1; PID:g56111  
R:Maggi, A.M.A.  
submitted to the EMBL Data Library, June 1991

A:Reference number: S16731  
A:Accession: S16731  
A:Molecule type: mRNA  
A:Residues: 1-487,'T',489-600 <MAG>  
A:Cross-references: EMBL:X61098; NID:g56120; PIDN:CAA43411.1; PID:g56121  
C:Comment: The steroid hormones and their receptors are involved in the regulation of eu  
C:Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly  
complex appears to recognize discrete DNA sequences upstream of transcriptional start sit  
C:Superfamily: estrogen receptor; erba transforming protein homology  
C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone recep  
F:1-184/Domain: amino-terminal <NH2>  
F:185-276/Domain: DNA binding #status predicted <DNA>  
F:188-461/Domain: erba transforming protein homology <ERBA>  
F:188-211/Region: zinc finger CCCC motif  
F:224-246/Region: zinc finger CCCC motif  
F:261-276/Region: nuclear location signal  
F:305-557/Domain: steroid binding #status predicted <STB>  
F:190,193,207,210/Binding site: zinc (Cys) #status predicted  
F:226,232,242,245/Binding site: zinc (Cys) #status predicted  
F:241,310/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 30.0%; Score 716.5; DB 1; Length 600;  
Best Local Similarity 39.2%; Pred. No. 5.7e-45;  
Matches 168; Conservative 82; Mismatches 128; Indels 51; Gaps 13;

QY 73 NGHONGLDSPPLY----PSAPITGSGSP---VRKLYDDCCSTIVEDPQTKE---YMLN 121  
Db 129 HGHQ-----VPYLLNEP\$AYAVRDTGPAPFYRSNSDNRQRNGRELS\$SSEKGMIMES 183  
QY 122 SMPKRLCLVCGDIASGYHYGVASCEACKAFEKRTIQNIEYSCPATNECEITKRRKSCQ 181  
Db 184 AKETRYCAVCNDYASGYHYGVWSCEGCKAFKRSIQGHNDYMCPATNQCTIDKRRKSCQ 243  
QY 182 ACRFMKCLKYGMKEGVRLLDRVRGGRQ-KYKR-RIDAEN-----SPYLNPOL 226  
Db 244 ACRLKCYEYGVMMKGIRKDR-RGGRMLKHKRQDDLEGRNEMGTSGDMRAANLWPSPLY 302  
QY 227 VQPAKK-----PYNKIVSHLLVAEPKTIYAMPDPTVPDSDIKALTTLCDLADRELVI 279  
Db 303 IKHTYKNSPALSLTADQWVSALLDAPPLIYSEYDPSRPFSEASMGLLTNLADRELVHM 362  
QY 280 IGWAKHIPGFSTLSLADQMSLQSAWMEIILGVYRSLSEFEDLVYADDYIMEDQSK- 338  
Db 363 INWAKRVGFGDLNLHDQVHLLCEAWLEIIMIGLVWSRMEHPGKLLFAPNLLDRNQSKC 422  
QY 339 LAGLLDLNNAITQLVKKYKSMKLEKEFVTLKATAIAN-----SDSMHIEDVEAVQX 390  
Db 423 VEGWEIFDMLLATSSRFRMNLQGEHFVCLKSIILNSGVYTFLSSTLKSLEKDHHR 482  
QY 391 LQDVLEALQDY--EAGQHMEDP--RRAKMLMTPLPLRQSTSTKAVOHFYNIKLEGVPMH 447  
Db 483 VLDKINDTLIHLMAKAGTLTQQQHRRLAQLLLILSHIRMSNKGMEHLYNMCKKNVPLY 542  
QY 448 KLFLEMLEA 456  
Db 543 DLLLEMLDA 551

RESULT 6  
147140  
estradiol receptor - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 20-Aug-1999  
C:Accession: I47140; S66250; S32402  
R:Bokenkamp, D.; Jungblut, P.W.; Thole, H.H.  
Mol. Cell. Endocrinol. 104, 163-172, 1994  
A:Title: The C-terminal half of the porcine estradiol receptor contains no post-translat  
A:Reference number: I47140; MUID:95080454; PMID:7988744  
A:Accession: I47140  
A:Molecule type: translated from GB/EMBL/DBD  
A:Status: preliminary;  
A:Molecule type: mRNA  
A:Residues: 1-595 <BOK>  
A:Cross-references: EMBL:Z37167; NID:g587554; PIDN:CAA85524.1; PID:g587555





```

A;Residues: 1-595 <GR2>
A;Cross-references: GB:X03635; NID:g31233; PIDN:CAA27284.1; PID:g31234
R;Ponglikitmongkol, M.; Green, S.; Chambon, P.
EMBO J. 7, 3385-3388, 1988
A;Title: Genomic organization of the human oestrogen receptor gene.
A;Reference number: A43021; MUID:89091079; PMID:3145193
A;Accession: A43021
A;Molecule type: DNA
A;Residues: 143-161;205-225;244-264;356-374,'G',376;402-422;447-460,'P',462-467;508-528
R;Keaveney, M.; Klug, J.; Gannon, F.
DNA Seq. 2, 347-358, 1992
A;Title: Sequence analysis of the 5' flanking region of the human estrogen receptor gene
A;Reference number: S27140; MUID:93075998; PMID:1476547
A;Accession: S27143
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-115 <KEA>
A;Cross-references: EMBL:X62462; NID:g31201; PIDN:CAA44322.1; PID:g31205
R;Pfeiffer, U.; Fecarotta, E.; Castagnetta, L.; Vidal, G.
Cancer Res. 53, 741-743, 1993
A;Title: Estrogen receptor variant messenger RNA lacking exon 4 in estrogen-responsive h
A;Reference number: S34000; MUID:93153765; PMID:791651
A;Accession: S34000
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 216-254,367-399,'G',401-434 <PFE>
A;Cross-references: EMBL:X73067; NID:g579865; PIDN:CAA51528.1; PID:g939886
R;Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.
Mol. Endocrinol. 6, 773-785, 1992
A;Title: Characterization of estrogen receptor variant mRNAs from human breast cancers.
A;Reference number: A41925; MUID:92293154; PMID:1603086
A;Accession: A41925
A;Molecule type: mRNA
A;Residues: 1-214,'ELPTLC' <DOT>
A;Cross-references: GB:M69297; NID:g182218; PIDN:AA58462.1; PID:g182219
A;Experimental source: clone 4; breast cancer
A;Note: sequence has been revised after extraction from NCBI backbone
A;Note: the complete sequence of neither the nucleotide nor the protein is shown in this
A;Note: sequence extracted from NCBI backbone (NCBIN:106580)
A;Accession: B41925
A;Status: significant sequence differences
A;Molecule type: mRNA
A;Cross-references: GB:M69296
A;Experimental source: clone 24; breast cancer
A;Note: sequence extracted from NCBI backbone (NCBIN:106597)
C;Comment: The steroid hormones and their receptors are involved in the regulation of eu
C;Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly
complex appears to recognize discrete DNA sequences upstream of transcriptional start sit
C;Genetics:
A;Gene: GDB:ESR1; ESR
A;Cross-references: GDB:119120; OMIM:133430
A;Map position: 6q25.1-6q25.1
A;Intons: 151/2; 215/1; 254/1; 366/1; 412/2; 457/1; 518/2
C;Superfamily: estrogen receptor; erbA transforming protein homology
C;Keywords: DNA binding; nucleus; phosphoprotein; steroid hormone receptor; transcriptio
F;1-120/Domain: amino-terminal <NH2>
F;121-299/Domain: DNA binding #status predicted <DNA>
F;183-456/Domain: erbA transforming protein homology <ERBA>
F;185-205/Region: zinc finger CCCC motif
F;221-245/Region: zinc finger CCCC motif
F;300-595/Domain: steroid binding #status predicted <STB>
F;236,305/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match          29.4%; Score 702; DB 1; Length 595;
Best Local Similarity 41.2%; Pred. No. 6.6e-44;
Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

QY      119 MLNSMPKRLCLVCGDIAGYHYGVASCEACKAFPKRTIOGNIYSCPATNECEITKRRRK 178
      |::||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       176 MESAKETRYCAVCNDYASGYHYGVWSCEGCKAFPKRSIQGHNDYMCPATNQCTIDKNRRK 235

QY      179 SCQACRFMKCLKVGMLEKGVRLDRVRGGRGQ-KYKR-RIDAEN-----SPYLN 223

```

```

Db      236  SCQACRLKRCYEVGMMKGIRKDR-RGGRMLKHKRQRDDGEGRGVEVSAGDMRAANLWPS      : : :
QY      224  POLVOPAKK-----PYNKIVSHLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADREL      : : :
Db      295  PLMTKRSKKNLSLALSLTADQWVSALLDAEPILYSEYDPTRFSEASMMGLLTNLADREL      : : :
QY      277  VWIIGWAKHIPGFSTLSIADQMSLQSAMMEIILIGVVYRSLSFEDELVYADDYIMEDQ      : : :
Db      355  YHMINWAKRVPGFVDLTLDQYHLLBCAWLEIMTGLVWRSMHPVKLLFAPNLLDRNQ      : : :
QY      337  SK-LAGLLDNNAILQLVKYKXSMKLEKEFVTLKAIALANS-----DSMH      : : :
Db      415  GKCEVGWEIIFDMLLATSPRFMMNLQGESEFVCLKSIILNLSGVYFTLSLTKSLEEKDH      : : :
QY      382  IEDVEAVQKLQDVLHEALQDYEAGQHMEDP-RRAGKMLMTPLLRQSTKAVQHFYNIKL      : : :
Db      475  IHRV--LDKITDTLIHLMA--KAGLTLQQQOHRLAQLLLSHIRHMSNKGMEHLYSMKC      : : :
QY      441  EGKVEPMHKLFLMEMLEA      456
Db      531  KNVVPDYDLLLEMLDA      546

```

## RESULT 9

## Q R X L E

estrogen receptor - African clawed frog

C:/Species: Xenopus laevis (African clawed frog)  
C:/Date: 28-Feb-1993 #20

```
C:/Date: 28-Feb-1992 #sequence_revision 14-Jul-1994 #text_change 15-Jun-1996
C:/Accession: A40907
```

R;weiller, I.J.; Lew, D.; Shapiro, D.T

Mol. Endocrinol. 1, 355-362, 1987

A;Title: The *Xenopus laevis* estrogen receptor: sequence homology with human and avian co

A;Reference number: A40907; MUID:90331927; PMID:3274894

A;Accession: A40907

A;Molecule type: mRNA

```
A;residues: 1-386 <WE1>
A:Cross-references: GB:1.20735
```

A/cross-references: GB:L20/35

C;Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly associated with corepressors and coactivators.

Complex appears to recognize discrete DNA sequences upstream of trans-

C: Superfamily: estrogen receptor; *erbA* transforming protein homology

**Keywords:** DNA binding; nucleus; phosphoprotein; steroid binding; E-1-174/Dexamethasone; anti-17 $\beta$ -oestradiol antibody

E:1775-266/Domains: DNA binding #status: unclassified

EF:178-448/Domain: erpA transforming protein homolog - erpA

F;178-201/Region: zinc finger CCCC motif <ERBA>

F;214-237/Region: zinc finger CCCC motif

F;251-266/Region: nuclear location signal

```
F;292-544/Domain: steroid binding #status predicted <STB>
```

E:180,183,197,200/Binding site: zinc (Cys)	#status predicted
E:216,222,225,227/Binding site: zinc (Cys)	1

```

E:231/Binding site: phosphate (con)
F:216,222,232,235/Binding site: zinc (Cys) #status predicted

```

residue	binding site	phosphonate (Ser)	(covalent)	#status	predicted
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9
10	10	10	10	10	10
11	11	11	11	11	11
12	12	12	12	12	12
13	13	13	13	13	13
14	14	14	14	14	14
15	15	15	15	15	15
16	16	16	16	16	16
17	17	17	17	17	17
18	18	18	18	18	18
19	19	19	19	19	19
20	20	20	20	20	20
21	21	21	21	21	21
22	22	22	22	22	22
23	23	23	23	23	23
24	24	24	24	24	24
25	25	25	25	25	25
26	26	26	26	26	26
27	27	27	27	27	27
28	28	28	28	28	28
29	29	29	29	29	29
30	30	30	30	30	30
31	31	31	31	31	31
32	32	32	32	32	32
33	33	33	33	33	33
34	34	34	34	34	34
35	35	35	35	35	35
36	36	36	36	36	36
37	37	37	37	37	37
38	38	38	38	38	38
39	39	39	39	39	39
40	40	40	40	40	40
41	41	41	41	41	41
42	42	42	42	42	42
43	43	43	43	43	43
44	44	44	44	44	44
45	45	45	45	45	45
46	46	46	46	46	46
47	47	47	47	47	47
48	48	48	48	48	48
49	49	49	49	49	49
50	50	50	50	50	50
51	51	51	51	51	51
52	52	52	52	52	52
53	53	53	53	53	53
54	54	54	54	54	54
55	55	55	55	55	55
56	56	56	56	56	56
57	57	57	57	57	57
58	58	58	58	58	58
59	59	59	59	59	59
60	60	60	60	60	60
61	61	61	61	61	61
62	62	62	62	62	62
63	63	63	63	63	63
64	64	64	64	64	64
65	65	65	65	65	65
66	66	66	66	66	66
67	67	67	67	67	67
68	68	68	68	68	68
69	69	69	69	69	69
70	70	70	70	70	70
71	71	71	71	71	71
72	72	72	72	72	72
73	73	73	73	73	73
74	74	74	74	74	74
7					

Query Match 29.1%: Score 694. DB 1, Length 886.

Best Local Similarity 34.5%; Pred. No. 2, 5e-43; Bengali 586;

Matches	161;	Conservative	86;	Mismatches	140;	Indels	80;	Gaps	12.
---------	------	--------------	-----	------------	------	--------	-----	------	-----

[illegible]

QY 43 PSSPA-----SLTDSVNHHS---PGSSDASGSY-----SSTMNGHQNGLD 80

[illegible]

22 PESPVEALKEQLSPFIHHGQVPEYYLESEQGTFAVREAPPTFYRSSSDNRQSGRE 158

81 SPPLYPAPILGGSGPVKLIYDNCSTIVEDPOTKCEWMI NSMBKPI CI YSCDTA SCVVVV 140

VANDEBROOK I VEDFQINCEIMENSMERKULCGLASGYHY 140

: : :

Db 159 R-----MSSANDKGPS-----MESTKETRYCAVCSDYASGYHY 192

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

141 GVASCEACKAFKRTIQGNI EYSCPATNECEITKRRKSCQACRFMKCLKVGMKEGVRJ 200

193 GWSCECCTAEFKBSCIOGINDUWANI TWO CMTNRNWTSTST : :

193 GVMSCEGKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGGRK 252

QY 201 DRVRCR-----OKYKBTDAEN----SPYTNBOLVODAKKRY NUTUIGU 241

-----XNLFVQFANNEY-----NKLVSHL 241

Db 253 DR-RGGRLLKHKRQKEEQKNDVDPSEIRTA SIWNP SVKSMKLSFVLSITAEQITSAI, 317

```

QY      242 LVASPEKIYAMPDPTVPDSDIKALTTLCDLADRELVIIGWAKHIPGSTLSLADQMSLL 301
      : || : : | | | | : : | | | | | | : | | | | | : | | | |
Db      312 MEAEAPIVYSEHDSTKPLSEASMTLLTNLADRELVHMIMWAKRVPGEVDLTTHDQVHLL 371
QY      302 QSAMWEILLGVVYRSLSEFEDLVYADDYIMDEDQSK-LAGLLDLNNAILQLVKKYKSMK 360
      : | | | | : : : | | : : | | : : | | : : | | : : : : : : : :
      : | | | | : : : | | : : | | : : | | : : | | : : : : : : : :
Db      372 ECAMLEILMWGLIWRSVHEHPKLSFAPNLLIDENQRCVEGLVEIFDMLVTTATRRFRMR 431
QY      361 LEKEEFVTLLKALIANSDSM-----HIEDVEAVQKLQDVLHEALQDYEAGQHM--E 409
      | | | | : : | | | | : : | | : : | | : : | | : : | | : :
      : | | | | : : | | | | : : | | : : | | : : | | : : | | : :
Db      432 LRGEEFICLKSIILNSGVYTFLSSTLESLEDTDLIHLIDKIIDTLVHFMKSGSLIQQ 491
QY      410 DPRRAGKMLMTPLLRQSTSTKAVQHFYNIKLEGPVPMHKLFLMLEA 456
      || : : : : : : | | : : | | : : | | : : | | : : | | : :
      : : : : : : | | : : | | : : | | : : | | : : | | : :
Db      492 QQRRLAQLLILSHIRHMSNKGMEHLYSMKCKNAVPLYDLLLEMLDA 538

RESULT 10
S58224
oestrogen receptor - Atlantic salmon (fragment)
C/Species: Salmo salar (Atlantic salmon)
C/Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Aug-1999
C/Accession: S58224
R/Rogers, S.A.; Llewellyn, L.; Ramsurn, V.P.; Sweeney, G.E.; Wigham, T.
submitted to the EMBL Data Library, July 1995
A/Reference number: S58224
A/Accession: S58224
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-535 <ROG>
A/Cross-references: EMBL:X89959; NID:g929901; PIDN:CAA61999.1; PID:g929902
C/Superfamily: estrogen receptor; erba transforming protein homology
C/Keywords: steroid hormone receptor; zinc finger
F;103-382/Domain: erba transforming protein homology <ERBA>

Query Match          28.2%; Score 674; DB 2; Length 535;
Best Local Similarity 32.8%; Pred. No. 6.7e-42;
Matches 161; Conservative 86; Mismatches 126; Indels 118; Gaps 15;

QY      38 FIKTEPSSPASLTDSVNHSPGSSDASGSY---SSTMNGHQGLDSPLYPSAPI--- 90
      | : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      26 FVSSSPQLSPFL-----HPGHGHPQSQSYLETST-----PLYRSSVVTNQ 68
QY      91 -----LGSQPYKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVC 131
      | | | | : : : : : : : : : : : : : : : : : : : : : :
Db      69 LSASEEKLCTSNRQOSYAAAGSGV--RVFEMANET-----RYCAVC 108
QY      132 GDIASGYHYGVASCEACKAFKKRTIQNTIEYSCPATNECITKRRRSQACRFMKCLKV 191
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      109 SDFASGYHYGFWSCEGCKAFFKRSIQGHNDYMCPATNQCTMDRNRKSCQACRLKCYEV 168
QY      192 GMLKEGVRLDRVGRG--QKYKR-----RIDAENSP----- 220
      || | | | | | | | | | | | | | | | | | | | | | | | |
      : | | | | : : | | | | : : | | : : | | : : | | : : | |
Db      169 GMYKGLRKD--RGRVLRKDKRKYCGPAGDREKPYGDLHRTAPPODGGRNSSSSSLSGG 226
QY      221 --YLNPOLVOPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELIV 278
      : : : : : : : : : : : : : : : : : : : : : : : : : :
      : | | | | : : | | | | : : | | : : | | : : | | : : | |
Db      227 GGWCGPRITMPPE---QVLFLQLQGAEPALCSROKVARPYTEVTMTLLTSMADKELVH 282
QY      279 IIGWAKHIPGSTLSLADQMSLLQSAMWEILLGVVYRSLSEFEDLVYADDYIMDEDQSK 338
      : | | | | | | | | | | | | | | | | | | | | | | | | | |
      : | | | | : : : | | : : | | : : | | : : | | : : : : : :
Db      283 MIAMAKKVPGFQELSLHDQVLLESSWLEVMIGLIWRSIHCPGKLIFAQDLLIDRSEGD 342
QY      339 -LAGLLDLNNAILQLVKKYKSMKLEKEFVTLLKALIANSDSM-----DSMHIEDVEA 387
      : | | : : : : : : : : : : | | | | | | | | | | : : | | : :
      : | | : : : : : : : : : : | | | | | | | | | | : : | | : :
Db      343 CVEGMABEIFDMLLATVSRFRMLKLKPEEFVCLKAILINSGAFSFCSNSVESLH--NSSA 400
QY      388 VQKLQDVLHEAL---QDYEAGQHMEDPRRAGKMLMTPLLRQSTSTKAVQHFYNIKLEGV 444
      | : : | | : : : : : : : : : : | | | | : : | | : : | | : :
      : : : | | : : : : : : : : : : | | | | : : | | : : | | : :
Db      401 VESMLDNITDALIHHSIGSGASVOQQPRROVQLLLLSHIRHMSNKGMEHLYSICKNKV 460
QY      445 PMHKLFLMLE 455

```

Db 461 PLYDLEMLD 471

RESULT 11  
T10423  
estrogen receptor - Oryzias sp. (strain d-rR)  
C/Species: Oryzias sp.  
A/Variety: strain d-rR  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Jun-2000  
C/Accession: T10423  
R/Okada, H.; Kawahara, T.; Yamashita, I.  
submitted to the EMBL Data Library, March 1994  
A/Description: Cloning of medaka estrogen receptor cDNA.  
A/Reference number: Z17013  
A/Accession: T10423  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-620 <OKA>  
A/Cross-references: EMBL:D28954  
A/Experimental source: strain d-rR, liver  
C/Genetics:  
A/Gene: MBR  
C/Superfamily: estrogen receptor; erba transforming protein homology  
C/Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation  
F;184-460/Domain: erba transforming protein homology <ERB>

Query Match	28.2%	Score 674	DB 2	Length 620
Best Local Similarity	35.2%	Pred. No. 8.1e-42		
Matches 161	Conservative 83	Mismatches 153	Indels 61	Gaps 11

  

DB	Query	Sequence	Score	DB	Length
Db	QY	43 PSSPASLTDSVNHSPGSSDASGSYSTMNGHONGLDSPPLVPSAPILGGSGPVKRLYD	102		
Db	QY	110 PLSPLVFWPSSPRLSP-----FMHPPSHHYLETSTPYRRSS---HQGASRE--D	154		
Db	QY	103 DCSSTIVEDPQTKCEYMLNS-----MPK--RLCLVCGDIASGYHYGVASCEACKAFKKR	154		
Db	QY	155 QCGSR--EDTCSLGLGAGAGAGFEMAKDRFCVAVCSDDYASGYHYGVWSCEGKAFKKR	212		
Db	QY	155 TIQGNIEYSCPATNECEITKRRKSCQACRFMKCLKVGMLEKEGVRLDRVRGGRQKKYKRI	214		
Db	QY	213 SIQGHNDYMCPATNQCTIDRNRKSCQACRLKCYEVGMKGGVAKDRIRILRDKRRRTG	272		
Db	QY	215 DAENBPYLN PQ-----LVQPAKKPYNKIVSHLLVAEPKITY	250		
Db	QY	273 VGDGKVKYKQGEHKTVHYDGRKRSSTGGGGGGGGRISVTSIPPEQVLLLLQGAEPILC	332		
Db	QY	251 AMPDPYVDSDIKALTTLCDLADRELVIIGAKHIPGFSTLSLADQMSLLQSAWMEILI	310		
Db	QY	333 SRQKLSRPYTEVTMTLLTSMADKELVHMTAWAKKLPGFLQLSTHDQVLLLESSWLEVLIM	392		
Db	QY	311 LGVYRISLSEFDELVYADDYIMDEDQSK-LAGLIDLNNAILQLVKKYKSMKLEKEEFVTL	369		
Db	QY	393 IGLIWRSHICPGKLI FAQDLILDRNEGCVEGMTEIFDMLLATASRFVRLKLPBEEFVCL	452		
Db	QY	370 KAIATLANSDSMH-----IEDVEAVQKIQDVLHEALQDY--EAGQ-HMEDPRRAGKML	418		
Db	QY	453 KAIITLNSGAFSFCGTGMEPLIHNSAAVQSMIDTITDALIHYSISQGYLAQEQARRQAOQLI	512		
Db	QY	419 MTLPLLRQTSKAVQHFYNIKLEGGKVPMHKLFLEMLEA	456		
Db	QY	513 LLLSHIRHMSNKGMEHLIYSMKCKKNKVPPLYDLEMLDA	550		

  

RESULT 12
A37197
estrogen receptor - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 20-Aug-1999
C;Accession: A37197; A40070
R;Pakdel, F.; Le Gac, F.; Le Goff, P.; Valotaire, Y.
Mol. Cell. Endocrinol. 71, 195-204, 1990
A;Title: Full-length sequence and in vitro expression of rainbow trout estrogen receptor



A;Reference number: A37197; MUID:91006824; PMID:2210031  
A;Accession: A37197  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-574 <PAK>  
A;Cross-references: GB:M31559  
R;Pakdel, F.; Le Guellec, C.; Vaillant, C.; Le Roux, M.G.; Valotaire, Y.  
Mol. Endocrinol. 3, 44-51, 1989  
A;Title: Identification and estrogen induction of two estrogen receptors (ER) messenger  
A;Reference number: A40070; MUID:89127284; PMID:2915648  
A;Accession: A40070  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 150-574 <PA2>  
A;Cross-references: GB:M31559; NID:g213783; PIDN:AAA49552.1; PID:g213784  
C;Superfamily: estrogen receptor; erba transforming protein homology  
C;Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; z  
F;145-421/Domain: erba transforming protein homology <ERBA>  
F;147-167/Region: zinc finger  
F;183-207/Region: zinc finger

Query Match 27.9%; Score 666.5; DB 2; Length 574;  
Best Local Similarity 35.0%; Pred. No. 2.6e-41;  
Matches 167; Conservative 79; Mismatches 142; Indels 89; Gaps 17;

QY 38 FIKTEPSSPASITDSVNHSPGSSDASGSY---SSTMNGHONGLDSPLYPSPAPILGG 93  
Db 64 FVSSSPQLSPQLSPFL--HPPSHHGLPSQSYLETST-----PLYRSSVTINQ 110  
QY 94 -SGPYRKLYDDCSSTIVEDPQ-----TKCEYMLNSMPKRLCLVCGDIASGYHYGA 143  
Db 111 LSAAEEL-----CIASDRQQSISAAGSVRVFEMANE--TRYCAVCSDFASGYHYGW 162  
QY 144 SCEACKAFPKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMKEGVRLDRV 203  
Db 163 SCEGCKAFPKRSIQGHNDYMCPTNQCTMDNRKRSQACRLKCYEVGMVKGRLKD-- 220  
QY 204 RGR--QYKRRRI---DAE-----NSPYLN-----POLVQPAKK 232  
Db 221 RGRVLRKDKRYCGPAGDREKPTVTWSTGQRPQDGRNSSSLNGGGWGRPRITMPEQ 280  
QY 233 PYNKIVSHLVAEPEKIYAMPDPTVPDSIDIKALTTLCDLADRELVIIGWAKHPIGFSTL 292  
Db 281 -----VLFLLQGQTPALCSROKVARPYTEVTMTLITSMADKELVHMIAWAKKYPGFQEL 335  
QY 293 SLADQMSLQSAWMEILLGVVYRSLSEFEDLVYADYIMDEDQSK-LAGLLDNNAILQ 351  
Db 336 SLHDQVQLLESSWLVMTGLIWRSHCPGKLIFAQDILIDRSEGDCEGMAEIEFDMLLA 395  
QY 352 LVKKYKSMKLEKEEFVTIKALALANS-----DSMHIEDVEAVOKLQDVLEHAL-- 399  
Db 396 TVSRFGMLKLPKEEFVCIKAILLNPGAFSCNSVESLH--NSSAVESMLDNTDALIH 453  
QY 400 -QDYEAGQHMEDPRRAGKMLMTPLRLRQSTKAVQHFYNIKLEKVPMEKLFLEMLE 455  
Db 454 HISHGASVQQPRQAQLLLLSHIRHMSNKGMEHLYSIKCKNKVPLYLLEMLD 510

RESULT 13

S64737  
80K estrogen receptor - human  
C;Species: Homo sapiens (man)  
C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Aug-1999  
C;Accession: S64737  
R;Pink, J.U.; Wu, S.Q.; Wolf, D.M.; Bilimoria, M.M.; Jordan, V.C.  
Nucleic Acids Res. 24, 962-969, 1996  
A;Title: A novel 80 kDa human estrogen receptor containing a duplication of exons 6 and  
A;Reference number: S64737; MUID:96174665; PMID:8600466  
A;Accession: S64737  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-701 <PIN>  
A;Cross-references: EMBL:U47678; NID:g1197854; PIDN:AAB00115.1; PID:g1197855

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996  
C;Superfamily: estrogen receptor; erba transforming protein homology  
C;Keywords: steroid hormone receptor; zinc finger  
F;183-456/Domain: erba transforming protein homology <ERBA>  
F;518-562/Domain: erba transforming protein homology #status atypical <ERB2>

Query Match 27.7%; Score 662; DB 2; Length 701;  
Best Local Similarity 30.7%; Pred. No. 7.3e-41;  
Matches 179; Conservative 80; Mismatches 132; Indels 192; Gaps 17;

QY 56 HSPGSSDASGSYSTMTNGHONGLDS--PPL---YPS-----APILGSGPVKRL 100  
Db 80 YGPGSEAAAFGS-----NGLGFFPLNSVSPSPMLLHPPQLSPFLQPHGQQVPY 130  
QY 101 Y--DDCSSTIVEDPQTKCEYMLNS-----MKRLCLVCGD 133  
Db 131 YLENPESGVTYREAGPPAFYRPNSDNRQGRERLASTNDKGSMAVESAKETRYCAVCND 190  
QY 134 IASGYHYGVAACEACKAFPKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGW 193  
Db 191 YASGYHYGVAWCEGCKAFPKRSIQGHNDYMCPTNQCTTIDKNRRKSCQACRLKCYEVGM 250  
QY 194 LKEGVRLDRVRGGRQ-KYKR-RIDAEN-----SPYLNQVLQVPAKK----- 232  
Db 251 MKGGLRKDR-RGRMLKHKRQDDGEGRGEVGSAGDMRAANLWPSPLMIKRSKNLSALS 309  
QY 233 -PYNKIVSHLVAEPEKIYAMPDPTVPDSIDIKALTTLCDLADRELVIIGWAKHPIGFST 291  
Db 310 LTAQVWSALDAEPPIYSEYDPTRFSEASWGLLTNLADRELVHMIAWAKRVGPFVD 369  
QY 292 LSLADQMSLQSAWMEILLGVVYRSLSEFEDLVYADYIMDEDQSK-LAGLLDNNAIL 350  
Db 370 LTLHDQVHLLECAWLEIMGLVWRSMHVPKLLFAPNLLDRNQCKVEGWEIEFDMLL 429  
QY 351 QLVKKYKSMKLEKEEFVTIKALALANS-----DSMH-----IED----- 384  
Db 430 ATSSRFMRMNLQGBEFVCLKSIILLNSGVYTFLLSTLKSLEEKDHIHRLDKITDTLTHL 489  
QY 385 -----VEAVQKLQDV----- 394  
Db 490 MAKAGLTLQOQHRLAQLLLLSHIRHMRNQCKVEGWEIEFDMLLATSSRFMRMNLQGE 549  
QY 395 -----LHEALQDYEAGQHM-----EDRR 413  
Db 550 EFVCLKSIILLNSGVYTFLLSTLKSLEEKDHIHRYLDKITDTLTHLMAKAGLTLQOQHRL 609  
QY 414 AGKMLMTPLRLRQSTKAVQHFYNIKLEKVPMEKLFLEMLEA 456  
Db 610 LAQLLLLSHIRHMSNKGMEHLYSMKCKNVPLYLLEMLDA 652

RESULT 14

JC5939  
estrogen receptor beta - human  
C;Species: Homo sapiens (man)  
C;Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 21-Jul-2000  
C;Accession: JC5939  
R;Ogawa, S.; Inoue, S.; Watanabe, T.; Hiroi, H.; Orimo, A.; Hosoi, T.; Ouchi, Y.; Muramat  
Biochem. Biophys. Res. Commun. 243, 122-126, 1998  
A;Title: The complete primary structure of human estrogen receptor beta (herbeta) and its  
A;Reference number: JC5939; MUID:98139878; PMID:9473491  
A;Accession: JC5939  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-530 <OGA>  
A;Cross-references: DDBJ:AB006590; NID:g2911151; PIDN:BAA24953.1; PID:g2911152  
C;Superfamily: estrogen receptor; erba transforming protein homology  
F;147-408/Domain: erba transforming protein homology <ERB>

Query Match 27.1%; Score 648; DB 2; Length 530;  
Best Local Similarity 35.1%; Pred. No. 5.5e-40;  
Matches 174; Conservative 93; Mismatches 145; Indels 84; Gaps 23;

QY	34	SCS-SFKTEPSS---PASTLTDVSNHSPGGS--SDASGYS--STMNGHONGLSDPELY	85
Db	17	NCSGSILPLEHGSITYIPSSYVDS--HHEYPAMTFYSPAVMNNYSIPSNVTNLEGGPGROTTS	75
QY	86	PSA--PILGSGPV---RKLVDGCSSTIVEDPQTK--CE--YMLNSMP-----	124
Db	76	PNVLMPTPGHLSPLVVRHQL-----SHLYAEPOKSPWCSEARSLEHTLPLVNRRETLKRVYSG	130
QY	125	-----KR---LCLVCGDIASGVHYGVASCEACKAFERKTIQGNIEYSCPATNE	169
Db	131	NRCA\$PVTGP\$SKRDAHFCAVCS\$DYASGYHYGVMSCEGCKAFERKSIQGHNDYICPATNQ	190
QY	170	CEITKRRRRKSCQACRFMKCLKVGMLKEGVRLDRVRGGROKYKRRIDAE-----	217
Db	191	CTIDKNRRKSCQACRLRKCYE\$VGMVKCGR--RERCGRYLVRQR\$AD\$EQLHCAGAKAKRS	248
QY	218	--NSPYLNQOLVQPAKKEPYNKIVSHLVAEPEKIYAMPDPTVPDS\$IKALTTLCOLADRE	275
Db	249	GGHAPRVR-ELLLDALSP-EQLVLTLL\$EAP\$PHV-LISRPA\$PTEASMMMSLT\$KLADKE	305
QY	276	L\$VYIGWAKHIPGFSTLSLADQMSLQ\$ANWEILLGVVYRSL\$FEDEL\$VYAD\$YIMDE\$	335
Db	306	L\$VHMISWAKKIPGFVELSL\$FDQVRLLES\$CNNEVLMMGLMMRSIDHPKLIIFAPDLVLRD	365
QY	336	Q\$K-LAGLLDLNNALLOLVK\$KYSKMKLEKEFVTLKAIALANDSMH-----IEDVEAVQ	389
Db	366	E\$K\$VEGIL\$IFDMLLATTSR\$FRELKLOHKEYL\$CVKAMILLNS-SMYPLVATQDADSSR	424
QY	390	KLQDVL\$EALQD-----YEAG-QHMEDPRRAGKMLMTLPLLRQTSY\$VQHFYNIKLEG	442
Db	425	KL\$HLLN-AVTDALVWYIAKSGISSQOQ\$MRLANLMLLSHV\$RHAS\$KNGEHLINMKCKN	483
QY	443	KYPMHKL\$LEMLEAKV	458
Db	484	V\$PVYD\$LLLEMLNAHV	499

```
QY      218 -----NSPYLNPOLVPAKKPYNKIIVSHLLVAEPEKIYAMPDPTVPDSDIK 263
           :: : : : | : : : | : : : | : : :
Db      184 EQLHCAGKAKRSGGHAPRY-ELLDDALSP-EQLVLTLLLEAPPHV-LISRPAPFTEAS 240

QY      264 ALTTLODLADRELVIITGMAKHIPGFSTLSLADOMSLOSAMEILLGVVYRSLSFEDE 323
           : : | | | : | : | | | : | : | : | : | : | : : : : :
Db      241 MMMSLTKLADKELVHMIWMAKKIPGVELSLFDQVRLLSCWMEMVMNGIMWRSIDHPGK 300

QY      324 LVYADDYIMDEDQSK-LAGILDLNNAIQLVKKYKSMKLEKEEFVTLLKAIALANSDSMH- 381
           || : | : | : | : | : : : : : : : : : | | | :
Db      301 LIFAPDLVLDRDEGCKVEGILETFMDMLLATTSRFRELKTLQHKEYLCVAKMILLNS-SMYP 359

QY      382 ----IEDVEAVOQLQDVLTLEALOD-----YEAG-QHMEDPRRAGKMILTLPRLROTSTK 430
           : : : : | : : : : : : : : : : : : : : : : : :
Db      360 LVTAQTQDADSSRKLAHLIN-AVTDALVWVIAKSGISQQQSMRLANLMLSSHVRHASNK 418

QY      431 AVQHFYNIKLEGKVPMHKLFTEMLEAKV 458
           : : | : | : | : | : | : | : | : | : | : | :
Db      419 GMEHILNMKCCKNVVPYVDLLEMTLANHV 446
```

```
Search completed: August 3, 2004, 12:42:16
Job time : 18 secs
```

RESULT 15  
S71400  
estrogen receptor beta - human  
C:Species: Homo sapiens (man)  
C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 20-Jun-2000  
C:Accession: S71400  
R:Mosselman, S.; Polman, J.; Dijkema, R.  
FEBS Lett. 392, 49-53, 1996  
A:Title: ER-beta: identification and characterization of a novel human estrogen receptor  
A:Reference number: S71400; MUID:96354875; PMID:8769313  
A:Accession: S71400  
A:Molecule type: mRNA  
A:Residues: 1-477 <MOS>  
A:Cross-references: EMBL:X99101; NID:g1518262; PIDN:CAA67555.1; PID:g1518263  
C:Superfamily: estrogen receptor; erba transforming protein homology  
C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone recep  
F;94-355/Domain: erba transforming protein homology <ERBA>  
F;96-116/Region: zinc finger CCCC motif  
F;132-156/Region: zinc finger CCCC motif  
F;167-182/Region: nuclear location signal  
F;52/Binding site: phosphate (Ser) (covalent) (by MAP kinase) #status predicted  
F;96,99,113,116/Binding site: zinc (Cys) #status predicted  
F;132,138,148,151/Binding site: zinc (Cys) #status predicted

	Query Match	27.0%;	Score 644.5;	DB 2;	Length 477;	
	Best Local Similarity	38.1%;	Pred. No. 8.6e-40;			
	Matches	Conservative	82;	Mismatches 117;	Indels 41;	Gaps 12;
QY	98 RKLYYDDCSSTIVEDPQTKECEYMLNSMPKRLCLVCGDIASGYHGVASCEACAKAFKKRTIQ	: : :     : : :		:	:	:
Dd	73 RKVSGNRCASPVTGPGSRD-----AHFCAYCSDSYASGYHGWSCSEGCKAFFKRSIQ	:		:	:	:
QY	158 GNIEYSCPATNECEITFKRRRKSQCACRFMKCLKYGMLEGEVRLDRVGRGQRKYKRRIADAE	: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :		:	:	:
Dd	126 GHNDYICPATNQCTIDKNRKSKCQACRLRKCYEVEGVKCGSR--RERCGYRLVRRRQRSAD	: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :		:	:	:

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:37:35 ; Search time 13 Seconds

(without alignments)  
1834.472 Million cell updates/sec

Title: US-10-054-841-4  
Perfect score: 2388  
Sequence: 1 MDSVELCLPESFSLHYEEL.....KLEGVPMHKLFILEMLEAKV 458

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2388	100.0	458	1	ERR3_HUMAN
2	1799	75.3	500	1	ERR2_HUMAN
3	1784	74.7	433	1	ERR2_RAT
4	1748	73.2	433	1	ERR2_MOUSE
5	1258.5	52.7	519	1	ERR1_MOUSE
6	1253.5	52.5	462	1	ERR1_MOUSE
7	716.5	30.0	600	1	ERR1_RAT
8	715	29.9	589	1	ERR1_CHICK
9	713.5	29.9	595	1	ERR1_PIG
10	712.5	29.8	587	1	ERR1_POEGU
11	711.5	29.8	595	1	ERR1_MESAU
12	710	29.7	599	1	ERR1_MOUSE
13	704	29.5	594	1	ERR1_HORSE
14	704	29.5	595	1	ERR1_HUMAN
15	694	29.1	586	1	ERR1_XENLA
16	690	28.9	569	1	ERR1_BRARE
17	687.5	28.8	579	1	ERR1_SPAU
18	684.5	28.7	585	1	ERR1_ORENI
19	682.5	28.6	622	1	ERR1_ONCMY
20	678	28.4	581	1	ERR1_PAGMA
21	676	28.3	554	1	ERR2_STUVU
22	674.5	28.2	617	1	ERR1_ICTPU
23	674	28.2	535	1	ERR1_SALISA
24	674	28.2	620	1	ERR1_ORYLA
25	671.5	28.1	472	1	ERR2_CHICK
26	666.5	27.9	472	1	ERR2_COTJA
27	661.5	27.7	583	1	ERR1_OREAU
28	661.5	27.7	673	1	ERR2_MICUN
29	655.5	27.4	530	1	ERR2_RAT
30	653.5	27.3	525	1	ERR1_MICUN
31	653	27.3	568	1	ERR2_ONCMY
32	652.5	27.3	527	1	ERR2_SHEEP
33	650	27.2	530	1	ERR2_MOUSE

34	648	27.1	530	1	ESR2_HUMAN	Q92731 homo sapien
35	648	27.1	568	1	ERB1_CARAU	Q9w669 carassius a
36	648	27.1	573	1	ESR2_ANGJA	Q13012 anguilla ja
37	642.5	26.9	527	1	ESR2_BOVIN	Q9xsb5 bos taurus
38	642.5	26.9	565	1	ESR3_MICUN	P57783 microgoni
39	640	26.8	526	1	ESR2_PIG	Q9xsw2 sus scrofa
40	639	26.8	557	1	ESR2_ORENI	Q9yh32 oreochromis
41	638.5	26.7	530	1	ESR2_CALJA	Q95171 callithrix
42	634	26.5	559	1	ESR2_SPAU	Q9w6m2 sparus aura
43	631	26.4	610	1	ERB2_CARAU	Q91a19 carassius a
44	629	26.3	575	1	ESR2_ICTPU	Q91a11 ictalurus p
45	623.5	26.1	441	1	RXRA_BRARE	Q90416 brachydanio

ALIGNMENTS

RESULT 1  
ERR3\_HUMAN  
ID ERR3\_HUMAN STANDARD; PRT; 458 AA.  
AC O75454; O96021; Q9R1F3;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Estrogen-related receptor gamma (Estrogen receptor related protein 3)  
DE (ERR gamma-2).  
GN ESRG OR NR3B3 OR ERG2 OR ERR3 OR KIAA0832.  
OS Homo sapiens (Human), and  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
CX NCBI\_TaxID=9606, 10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM LONG).  
RC SPECIES=Human; TISSUE=Brain;  
RX MEDLINE=99173874; PubMed=10072763;  
RA Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,  
RA Liu Q., Caskey C.T.;  
RT "Identification of two hERR2-related novel nuclear receptors utilizing  
RT bioinformatics and inverse PCR.";  
RL Gene 228:101-109(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM LONG).  
RC SPECIES=Human; TISSUE=Brain;  
RX MEDLINE=99156230; PubMed=10048485;  
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 5:355-364(1998).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
RC SPECIES=Human;  
RX MEDLINE=98341123; PubMed=9676434;  
RA Eudy J.D., Yao S.F., Weston M.D., Ma-Edmonds M., Talmadge C.B.,  
RA Cheng J.J., Kimberling W.J., Sumegi J.;  
RT "Isolation of a gene encoding a novel member of the nuclear receptor  
RT superfamily from the critical region of Usher syndrome type IIa at  
RT 1q41.";  
RL Genomics 50:382-384(1998).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Mouse;  
RX MEDLINE=99357798; PubMed=10428842;  
RA Hong H., Yang L., Stallcup M.R.;  
RT "Hormone-independent transcriptional activation and coactivator  
RT binding by novel orphan nuclear receptor ERR3.";  
RL J. Biol. Chem. 274:22618-22626(1999).  
CC -!- FUNCTION: ORPHAN RECEPTOR. BINDS SPECIFICALLY TO AN ESTROGEN  
CC RESPONSE ELEMENT AND ACTIVATES REPORTER GENES CONTROLLED BY  
CC ESTROGEN RESPONSE ELEMENTS.  
CC -!- SUBUNIT: Homodimer (Potential).



-1- SUBCELLULAR LOCATION: Nuclear (potential).  
-1- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=Long;  
IsoId=075454-1; Sequence=Displayed;  
Name=Short;  
IsoId=075454-2; Sequence=VSP\_003702;  
Note=No experimental confirmation available;  
-1- TISSUE SPECIFICITY: EXPRESSED IN THE BRAIN, LUNG, BONE MARROW,  
ADRENAL GLAND, TRACHEA, SPINAL CORD AND THYROID GLAND.  
-1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN FETAL BRAIN AND  
ALSO IN THE FETAL KIDNEY, LUNG AND LIVER.  
-1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3  
subfamily.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF094518; AAC99410.1; -;  
EMBL; AB020639; BAA74855.1; -;  
EMBL; AF058291; AAC39899.1; -;  
EMBL; AF117254; AAD48369.1; -;  
PDB; 1KV6; 25-JAN-03.  
TRANSFAC; T02766; -;  
TRANSFAC; T04644; -;  
TRANSFAC; T04645; -;  
TRANSFAC; T04648; -;  
Gene; HGNC:3474; ESRRG.  
MIM; 602969; -;  
MGD; MGI:1347056; Esrrg.  
DR GO; GO:0005634; C:nucleus; NAS.  
DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; NAS.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.  
DR InterPro; IPR005336; Hormone\_rec\_lig.  
DR InterPro; IPR001723; Strhmn\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR PRODom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
Zinc-finger; Activator; Alternative splicing; 3D-structure.  
FT DNA\_BIND 128 193 NUCLEAR RECEPTOR-TYPE.  
FT ZN\_FING 128 148 C4-TYPE.  
FT ZN\_FING 164 188 C4-TYPE.  
FT VARSPPLIC 1 23 Missing (in isoform Short).  
FT 151 151 F -> S (IN REF. 3).  
FT CONFLICT 155 155 T -> K (IN REF. 3).  
FT CONFLICT 158 158 G -> A (IN REF. 3).  
FT CONFLICT 271 271 L -> C (IN REF. 3).  
FT CONFLICT 313 313 V -> F (IN REF. 3).  
FT CONFLICT 458 458 V -> VC (IN REF. 3).  
SQ SEQUENCE 458 AA; 51305 MW; 63D36CFD37573152 CRC64;

Query Match 100.0%; Score 2388; DB 1; Length 458;  
Best Local Similarity 100.0%; Pred. No. 2e-168;  
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEEELLCRMSNKRHRIDSSCSSEFIKTEPSSPASITDSVNHSPGG 60  
Db 1 MDSVELCLPESFSLHYEEELLCRMSNKRHRIDSSCSSEFIKTEPSSPASITDSVNHSPGG 60

QY 61 SSDASGSYSTNMGNHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYML 120  
Db 61 SSDASGSYSTNMGNHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTCEYML 120  
QY 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIEYSCPATNECEITKRRKSC 180  
Db 121 NSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIEYSCPATNECEITKRRKSC 180  
QY 181 QACRFMKCLKVGMKEGYRLDRVRGGRQYKRRIDAENSPLYNPQLVQPAKKPYNKIVSH 240  
Db 181 QACRFMKCLKVGMKEGYRLDRVRGGRQYKRRIDAENSPLYNPQLVQPAKKPYNKIVSH 240  
QY 241 LLVAPEKIYAMPDPYVDSIDIKALTTLCDLADRELVIIGWAKHIPGFSTLSDQMSL 300  
Db 241 LLVAPEKIYAMPDPYVDSIDIKALTTLCDLADRELVIIGWAKHIPGFSTLSDQMSL 300  
QY 301 LOSAMMEILILGVVYRSLFDEELVYADDYIMDEDQSKLAGLLDNNAILQLVKKYKSMK 360  
Db 301 LOSAMMEILILGVVYRSLFDEELVYADDYIMDEDQSKLAGLLDNNAILQLVKKYKSMK 360  
QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLQDVILHEALQDYEAGQHMEDPRRAGKMLMT 420  
Db 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLQDVILHEALQDYEAGQHMEDPRRAGKMLMT 420

RESULT 2

ERR2\_HUMAN STANDARD; PRT; 500 AA.

AC 095718; Q9HCB4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Steroid hormone receptor ERR2 (Estrogen-related receptor, beta)  
DE (ERR-beta) (Estrogen receptor-like 2) (ERR beta-2).  
GN ESRRB OR NR3B2 OR ESRB2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=99173874; PubMed=10072763;  
RA Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,  
RA Liu Q., Caskey C.T.;  
RT "Identification of two hERR2-related novel nuclear receptors utilizing  
bioinformatics and inverse PCR.";  
RL Gene 228:101-109(1999).  
RN [2]  
RP SEQUENCE OF 264-500 FROM N.A.  
RA Rowen L., Madan A., Qin S., Baradarani L., Birditt B., Bloom S.,  
RA Dors M., Dickhoff R., Fleetwood P., Harrison G., Kaur A., Madan A.,  
RA Nesbitt R., Traicoff R., Hood L.;  
RT "Sequencing of human chromosome 14q24.3 region.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).  
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3  
subfamily.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF094517; AAC99409.1; -;

```
DR EMBL; AC008050; AAG17472.1; -.
DR HSSP; P03372; 1HCQ.
DR TRANSFAC; T02766; -.
DR Genew; HGNC:3473; ESRRB.
DR MIM; 602167; -.
DR GO; GO:0004879; F.ligand-dependent nuclear receptor activity; TAS.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001723; Stdhmrn_receptor.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR Prodom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
FT DNA_BIND 103 168 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 103 123 C4-TYPE.
FT ZN_FING 139 163 C4-TYPE.
SQ SEQUENCE 500 AA; 55619 MW; C219C84D914DF1C6 CRC64;
```

Query Match 75.3%; Score 1799; DB 1; Length 500;  
Best Local Similarity 78.1%; Pred. No. 5.5e-125;  
Matches 339; Conservative 50; Mismatches 43; Indels 2; Gaps 2;

```
QY 24 MSNKRHRIDSSCSSFIKTEBPSSPASLTDSVNHSPGSSSDASGSYSTMNGHQGLDSP 83
Db 1 MSSEDRHLGSSCGSFIKTEBPSSSGIDALSHSPSGSSDASGFGMALGTHANGLDSP 60

QY 84 LYPSAPILGSGSPVRKLYDCSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143
Db 61 MFAGAG-LGGT-PCRKSYEDCAGIMEDSAIKCEYMLNAIPKRLCLVCGDIASGYHYGVA 118

QY 144 SCEACKAFKRTIQNIEYSCPATNECEITKRRRKSQAACRFMKCLKVGMKEGVRLDRV 203
Db 119 SCEACKAFKRTIQNIEYSCPATNECEITKRRRKSQAACRFMKCLKVGMKEGVRLDRV 178

QY 204 RGGROKYKRRIDAENSPYLNQVOPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSIK 263
Db 179 RGGROKYKRRILDSESSPYLSIQISPPAKKPLTKIVSYLLVAEPDKLYAMPPPGMDIK 238

QY 264 ALTTLCDLADRELVIIGWAKHIPGSTLSLADQMSLIQSAWMEILILGVVRSLSFEDE 323
Db 239 ALTTLCDLADRELVIIGWAKHIPGFSSLSLGDQMSLIQSAWMEILILGIVRSLSFYDDK 298

QY 324 LYYADDYIMDEDQSKLAGLIDNNAILQLVKYKSMKLEKEEFVTLKATALANDSMHIE 383
Db 299 LYYAEDYIMDEHSLRGLLELYRAILQLVRRYKYLKVEKEEFVTLKATALANDSMYIE 358

QY 384 DVEAVQKLQDVLHEALQDYEAQGHEMDPRRAGKMLMTLLRLROTSTKAVQHFYNIKEGK 443
Db 359 DLEAVQKLQDVLHEALQDYELSQLRHEEPWRTGKLLTLPLLRQTAAKAVQHFYSVKLQKG 418

QY 444 VPMHKLFLFMLBAK 457
Db 419 VPMHKLFLMLEAK 432
```

RESULT 3  
ERR2\_RAT  
ID ERR2\_RAT STANDARD; PRT; 433 AA.  
AC P11475;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Steroid hormone receptor ERR2 (Estrogen-related receptor, beta) (ERR-beta)  
DE beta) (Estrogen receptor-like 2).  
GN ESRRB OR NR3B2 OR ESRL2 OR ERR2.  
OS Rattus norvegicus (Rat).

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=88122546; PubMed=3267207;
RA Giguere V., Yang N., Segui P., Evans R.M.;
RT "Identification of a new class of steroid hormone receptors.";
RL Nature 331:91-94(1988).
RN [2]
RP SHOWS THAT REF.1 ORIGINATES FROM RAT.
RX MEDLINE=99173874; PubMed=10072763;
RA Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,
RA Liu Q., Caskey C.T.;
RT "Identification of two hERR2-related novel nuclear receptors utilizing
RT bioinformatics and inverse PCR.";
RL Gene 228:101-109(1999).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC -!- CAUTION: Was originally (Ref.1) thought to originate from human
CC but was later shown (Ref.2) to be derived from rat.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; X51417; CA35779.1; -.
DR HSSP; P03372; 1HCQ.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001723; Stdhmrn_receptor.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR Prodom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
FT DNA_BIND 103 168 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 103 123 C4-TYPE.
FT ZN_FING 139 163 C4-TYPE.
SQ SEQUENCE 433 AA; 48287 MW; EE2C4C5B2F9A3E13 CRC64;
```

Query Match 74.7%; Score 1784; DB 1; Length 433;  
Best Local Similarity 76.6%; Pred. No. 5.7e-124;  
Matches 333; Conservative 51; Mismatches 49; Indels 2; Gaps 1;

```
QY 24 MSNKRHRIDSSCSSFIKTEBPSSPASLTDSVNHSPGSSSDASGSYSTMNGHQGLDSP 83
Db 1 MSSEDRHLGSSCGSFIKTEBPSSSGIDALSHSPSGSSDASGFGMALGTHANGLDSP 60

QY 84 LYPSAPILGSGSPVRKLYDCSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143
Db 61 MFAGAGL-GGNPCRKSYEDCTSGIMEDSAIKCEYMLNAIPKRLCLVCGDIASGYHYGVA 118

QY 144 SCEACKAFKRTIQNIEYSCPATNECEITKRRRKSQAACRFMKCLKVGMKEGVRLDRV 203
Db 119 SCEACKAFKRTIQNIEYSCPATNECEITKRRRKSQAACRFMKCLKVGMKEGVRLDRV 178

QY 204 RGGROKYKRRIDAENSPYLNQVOPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSIK 263
Db 179 RGGROKYKRRILDSESSPYLSIQISPPAKKPLTKIVSYLLVAEPDKLYAMPPDDVPEGDIK 238
```





RP FUNCTION.  
RX MEDLINE=97415618; PubMed=9271417;  
RA Sladek R., Bader J.-A., Giguere V.;  
RT "The orphan nuclear receptor estrogen-related receptor alpha is a  
RT transcriptional regulator of the human medium-chain acyl coenzyme A  
RT dehydrogenase gene.";  
RL Mol. Cell. Biol. 17:5400-5409(1997).  
CC -1- FUNCTION: BINDS TO AN ERR-ALPHA RESPONSE ELEMENT (ERRE) CONTAINING  
CC A SINGLE CONSENSUS HALF-SITE, 5'-TNAAGTCA-3'. CAN BIND TO THE  
CC MEDIUM-CHAIN ACYL COENZYME A DEHYDROGENASE (MCAD) RESPONSE ELEMENT  
CC NRRE-1 AND MAY ACT AS AN IMPORTANT REGULATOR OF MCAD PROMOTER.  
CC BINDS TO THE C1 REGION OF THE LACTOFERRIN GENE PROMOTER.  
CC -1- SUBUNIT: Binds DNA as a monomer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X51416; CAA35778.1; -.  
DR EMBL; L38487; AAB17015.1; -.  
DR PIR; A29345; A29345.  
DR HSSP; P03372; 1HCQ.  
DR TRANSPAC; T02765; -.  
DR Genew; HGNC:3471; ESRRA.  
DR MIM; 601998; -.  
DR GO; GO:0005634; C:nucleus; TAS.  
DR GO; GO:0003677; F:DNA binding; TAS.  
DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; TAS.  
DR InterPro; IPR000536; Hormone\_rec\_lig.  
DR InterPro; IPR001723; Stdrhm\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHOMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOI1; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger.  
FT DNA BIND 176 241 NUCLEAR RECEPTOR-TYPE.  
FT ZN FING 176 196 C4-TYPE.  
FT ZN FING 212 236 C4-TYPE.  
FT CONFLICT 39 44 ATSSRR -> SYKOPA (IN REF. 2).  
FT CONFLICT 288 288 A -> AA (IN REF. 1).  
FT CONFLICT 441 450 AEAVEQLREA -> EPRIMSSCEKL (IN REF. 1).  
SQ SEQUENCE 519 AA; 55439 MW; FA30F3ECTBF7B77F CRC64;  
Query Match 52.7%; Score 1258.5; DB 1; Length 519;  
Best local Similarity 55.1%; Pred. No. 3.7e-85;  
Matches 256; Conservative 67; Mismatches 79; Indels 63; Gaps 8;  
QY 22 CRMSNKRHIIDSSCSS-----FIKTEBSPASLTDSVNHSPGSSD-----A 64  
DB 85 CPTALPEPQYTSAMSSQVGVIEPLYIKAEFASP-----DSPKGSSETEPEPVALA 135  
QY 65 SG-SYSTMGHQNGLDSPPLYPAPILGSGFVRKLYDDCSSTIVEDPQTKCEYMLNSM 123  
DB 136 PGAPATRCLEPGHKEEEDGE-----GAGP-----GEGGGKLVLSL 171  
QY 124 PKRLCLVCGDIASGYHYGVASCEACKAFKRTIQNIETSCPATNECETTKRRKSCQAC 183  
DB 172 PKRLCLVCGDVASGYHYGVASCEACKAFKRTIQSIETSCPASNECETTKRRKACQAC 231

QY 184 RFMKCLKVGMKEGVRIDRVGRGRQKYKRIDAENSPYLNQLVQP-----AKKEYN 235  
DB 232 RFTKCLRVGMLKEGVRLDRVRGRGRQYKRRPEVDPLPPGPPFAGPLAVAGPRKTA PVN 291  
QY 236 KIVSHLVAEPEKIYAMPDPTVPDSIDKALTLLCDLADRELVTITGMAKHIPFSTLSLA 295  
DB 292 ALVSHLVAEPEKIYAMPDPAGPDGHLPAVATLCDLPREIIVTISWAKSIPGFSSLSLS 351  
QY 296 DQMSLQSANMEILILGVYRSLSEFEDLVYADDYIMDEDQSKLAGLIDNNAIIQLVK 355  
DB 352 DQMSVLQSVNMEVLVLGVAQRSLPLQDELAFAEDLVLDEEGARAAGLIGALILQLVRR 411  
QY 356 YKSMKLEKEEFVTLKATALANSDSMHIEDVEAVQKLQDVLHEALQDYEAQQ---HMEDP 411  
DB 412 LQALRLEREVEYVLKALALANSDSVHIEDAVEQLREALHEALLEYEAGRAGPGGAER 471  
QY 412 RRAGKMLMTLPRLRQTSTKAVQHFYNIKLEGGKVPMHKLFLEMLFA 456  
DB 472 RRAGRLTLTLPRLRQTAKVLAFHYGVKLEGGKVPMHKLFLEMLEA 516  
RESULT 6  
ERR1\_MOUSE  
ID ERR1\_MOUSE STANDARD; PRT; 462 AA.  
AC 008580;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Steroid hormone receptor ERR1 (Estrogen-related receptor, alpha)  
DE (ERR-alpha) (Estrogen receptor-like 1) (Fragment).  
GN ESRRA OR NR3B1 OR ERR1 OR ESTRRA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C;  
RX MEDLINE=97415618; PubMed=9271417;  
RA Sladek R., Bader J.-A., Giguere V.;  
RT "The orphan nuclear receptor estrogen-related receptor alpha is a  
RT transcriptional regulator of the human medium-chain acyl coenzyme A  
RT dehydrogenase gene.";  
RL Mol. Cell. Biol. 17:5400-5409(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain, and Kidney;  
RX MEDLINE=98121983; PubMed=9460651;  
RA Shigeta H., Zuo W., Yang N., DiAugustine R., Teng C.T.;  
RT "The mouse estrogen receptor-related orphan receptor alpha 1:  
RT molecular cloning and estrogen responsiveness.";  
RL J. Mol. Endocrinol. 19:299-309(1997).  
CC -1- FUNCTION: BINDS TO AN ERR-ALPHA RESPONSE ELEMENT (ERRE) CONTAINING  
CC A SINGLE CONSENSUS HALF-SITE, 5'-TNAAGTCA-3'. CAN BIND TO THE  
CC MEDIUM-CHAIN ACYL COENZYME A DEHYDROGENASE (MCAD) RESPONSE ELEMENT  
CC NRRE-1 AND MAY ACT AS AN IMPORTANT REGULATOR OF MCAD PROMOTER. MAY  
CC FUNCTION AS A MODULATOR OF THE ESTROGEN SIGNALING PATHWAY IN THE  
CC UTERUS.  
CC -1- SUBUNIT: Binds DNA as a monomer (Probable).  
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -1- TISSUE SPECIFICITY: MOST HIGHLY EXPRESSED IN KIDNEY, HEART, AND  
CC BROWN ADIPOCYTES. ALSO FOUND IN UTERUS, CERVIX AND VAGINA.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN AN ORGAN SPECIFIC MANNER THROUGH  
CC MID-TO LATE EMBRYONIC DEVELOPMENT WITH PERSISTENT HIGH-LEVEL  
CC EXPRESSION IN BROWN ADIPOSE TISSUE AND INTESTINAL MUCOSA.  
CC -1- INDUCTION: ACTIVATED BY DIETHYLSTILBESTROL (DES) AND ESTRADIOL IN  
CC THE UTERUS.  
CC -1- PTM: Phosphorylated (Probable).  
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger; Steroid-binding; Phosphorylation.  
FT DOMAIN 1 189 MODULATING.  
FT DNA\_BIND 190 255 NUCLEAR RECEPTOR-TYPE.  
FT ZN\_FING 190 210 C4-TYPE.  
FT ZN\_FING 226 250 C4-TYPE.  
FT DOMAIN 256 315 HINGE.  
FT DOMAIN 316 600 STEROID-BINDING.  
FT DOMAIN 64 71 POLY-ALA.  
FT DOMAIN 171 174 POLY-SER.  
FT MOD\_RES 109 109 PHOSPHORYLATION (BY CDK2) (BY  
FT MOD\_RES 111 111 SIMILARITY).  
FT MOD\_RES 111 111 PHOSPHORYLATION (BY CDK2) (BY  
FT MOD\_RES 123 123 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 172 172 PHOSPHORYLATION (BY CK2) (BY SIMILARITY).  
FT MOD\_RES 542 542 PHOSPHORYLATION (BY TYR-KINASES) (BY  
FT CONFLICT 488 488 N -> T (IN REF. 3).  
SQ SEQUENCE 600 AA; 67030 MW; C9C7D8CACE0F57D8 CRC64;

Query Match 30.0%; Score 716.5; DB 1; Length 600;  
Best Local Similarity 39.2%; Pred. No. 3.7e-45;  
Matches 168; Conservative 82; Mismatches 128; Indels 51; Gaps 13;

QY 73 NGHONGLDSPPLY---PSAPILGSGP---VRKLYDDCSSTIVEDPTQKCE---YMLN 121  
Db 129 HGHQ-----VPYYLENESAAYAVRDTGPAPFYRSNSDNRRQNGRERLSSSEKGNMIMES 183  
QY 122 SMPRLCLVCGDIASGYHYGVASCEACKAFKRTIOGNIIEYSCPATNECEITKRRKSCQ 181  
Db 184 AKETRYCAVCNDYASGYHYGVASCEGCKAFKRSIQGHNDYMCPTATNQCTIDKRRKSCQ 243  
QY 182 ACRFMKCLVGMKKEGVRLDRVGRGRQ-KYKR-RIDAEN-----SPYLNQOL 226  
Db 244 ACRLRKCYEVGMKKGIRKDR-RGGRMLKHKRQDDLEGRNEMGTSGDMRANULMPSPLV 302  
QY 227 VQPAKK-----PYNKIVSHLVAEPEKIYAMPDPTVPDSIDIKALTTLCULADRELVI 279  
Db 303 IKHTKNSPALSTADQVVSALLDAEPPLIYSEYDPSRPFSEASMGILLTNLADRELVHM 362  
QY 280 IGWAKHIPGFSTLSLADQMSLQSAWMEILLGVYRSLSFEDELVYADYIMDEDQSK- 338  
Db 363 INWAKRVPGFDLNLHDQVHLBCAWLEIMGLVWRSMHPGKLPAPNLLIDRNOGKC 422  
QY 339 LAGLLDLNNAILLQVKKYSMKLEKEFEVTLKALALAN-----SDSMHIEDVEAVOK 390  
Db 423 VEGMVEIFDMILLATSSRFRMNLQGEFVCLKSIIILNSGVYTFLSSTLKSLEKDHHR 482  
QY 391 LQDVLEALQDY--EAGQHMEDP-RRAGKMLMTPLRLQSTSTKAVQHFYNIKLEKVPMH 447  
Db 483 VLDKINDTLTHMAKAGLTLQQOHRRLAQLLLISHIRHMSNKGMEHLNMYKCKKNVPLV 542  
QY 448 KLFLEMLEA 456  
Db 543 DLLEMLDA 551

RESULT 8  
ESR1\_CHICK  
ID ESR1\_CHICK STANDARD; PRT; 589 AA.  
AC P06212;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).  
GN ESR1 OR NR3A1 OR ESR.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=86247578; PubMed=3755102;  
RA Krust A., Green S., Argos P., Kumar V., Walter P., Bornert J.-M.,  
RA Chambon P.;  
RT "The chicken oestrogen receptor sequence: homology with v-verba and  
RT the human oestrogen and glucocorticoid receptors.";  
RL EMO J. 5:891-897(1986).  
CC -!- FUNCTION: The steroid hormones and their receptors are involved in  
CC the regulation and differentiation in target tissues.  
CC -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with ER-  
CC beta.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,  
CC a DNA-binding domain and a C-terminal steroid-binding domain.  
CC -!- MISCELLANEOUS: In the absence of ligand, steroid hormone receptors  
CC are thought to be weakly associated with nuclear components;  
CC hormone binding greatly increases receptor affinity. The hormone-  
CC receptor complex appears to recognize discrete DNA sequences  
CC upstream of transcriptional start sites.  
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3  
CC subfamily.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; X03805; CAA27433.1; -.  
DR PIR; A40914; QRCHE.  
DR HSSP; P03372; IHCP.  
DR TRANSFAC; T00264; -.  
DR GO; GO:0016585; C:chromatin remodeling complex; ISS.  
DR GO; GO:0016020; C:membrane; ISS.  
DR GO; GO:0030284; F:estrogen receptor activity; ISS.  
DR GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.  
DR GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.  
DR InterPro; IPR000536; Hormone\_rec\_1ig.  
DR InterPro; IPR001292; Oestrgn\_receptor.  
DR InterPro; IPR001723; Stdrhm\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF02159; Oest\_recep; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHOMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger; Steroid-binding.  
FT DOMAIN 1 178 MODULATING.  
FT DNA\_BIND 179 244 NUCLEAR RECEPTOR-TYPE.  
FT ZN\_FING 179 199 C4-TYPE.  
FT ZN\_FING 215 239 C4-TYPE.  
FT DOMAIN 245 304 HINGE.  
FT DOMAIN 305 589 STEROID-BINDING.  
SQ SEQUENCE 589 AA; 66746 MW; 1B092233C770A54B CRC64;

Query Match 29.9%; Score 715; DB 1; Length 589;  
Best Local Similarity 36.2%; Pred. No. 4.7e-45;  
Matches 167; Conservative 90; Mismatches 138; Indels 66; Gaps 12;

QY 38 FIKTEPSSPASLTDSVNHHS---PGSSDASGSYSTMTNGHONGLDSPPLY--PSAPILG 92  
Db 104 FLQTA-----QLSPFIHHHSQVPPYYLENEGGSF-----GMRBA--APPAFYRPS----- 148



QY 93 GSGPVRKLYDDCSSTIVEDPQTCSE---YMLNSMPKRLCLVCGDIASGYHYGASCEAC 148  
Db 149 -----DNRHRSIRERMSSTNEKSGLSMESTKETRYCAVCNDYASGYHYGWSCEGC 199  
QY 149 KAFKKRTIQGNIYESCPAINECEITKRRRKSQAACRFMKCLKVGMLEKGVRLDRVRGRQ 208  
Db 200 KAFKKRSIQGHNDYMCPATNQCTIDKNRRKSQAACRLRKCYEVGMKGIKRRDRGGEMM 259  
QY 209 KYKRRIDAENSP-----YLNPLQVOPAKK-----PYNKIVSHLLVAEPE 247  
Db 260 KOKRQREQDSRNGEASSTELRAPLTWTSPLVVKHNKNKSPALSLTAEQVVSALLAEAPP 319  
QY 248 KIYAMPDPTVPDSDIKALTTLCIDLADRELVTIGWAKHIPGFSTLSLADQMSLQSAWME 307  
Db 320 IVYSEYDPNRPENASMTLLTNLADRELVHMIMWAKRVPGFVDLTLDHQQVHLLCAWLE 379  
QY 308 ILLGVYKSLSFEDLVYADDYIMDEDQSK-LAGLLDLNNAIQLVKKYKSMKLEKEEF 366  
Db 380 ILMIGLVKSMHEHPGKLLFAPNLLDRNGCKVEGMBEIFDMLLATFARFRMNNLQGEFF 439  
QY 367 VTLKAIALAN-----SDSMHIEDVEAVOKLQDVLEHALQDYEAQOHM---EDPRRAG 415  
Db 440 VCLKSIILLNSGVYTFLLSTLKSLEERDYIHRVLDKITDTLIHLMAKSGLSLQQQHRRLA 499  
QY 416 KMLMTPLLRQSTTKAVQHFYNIKLEKVPMEKLFLEMLEA 456  
Db 500 QLLILSHIRHMSNMKGMEHLYNMCKXNVPLYDLLEMLDA 540

RESULT 9  
ESR1\_PIG STANDARD; PRT; 595 AA.  
ID ESR1\_PIG  
AC Q29040;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).  
GN ESR1 OR NR3A1 OR ESR.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Uterus;  
RX MEDLINE=95080454; PubMed=7988744;  
RA Boekenkamp D., Jungblut P.W., Thole H.H.;  
RT "The C-terminal half of the porcine estradiol receptor contains no  
RT post-translational modification: determination of the primary  
RT structure."  
RL Mol. Cell. Endocrinol. 104:163-172(1994).  
CC -|- FUNCTION: Nuclear hormone receptor. The steroid hormones and their  
CC receptors are involved in the regulation of eukaryotic gene  
CC expression and affect cellular proliferation and differentiation  
CC in target tissues (By similarity).  
CC -|- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with  
CC ESR2. Interacts with NCOA3, NCOA5 and NCOA6 coactivators, leading  
CC to a strong increase of transcription of target genes (By  
CC similarity).  
CC -|- SUBCELLULAR LOCATION: Nuclear.  
CC -|- DOMAIN: Composed of three domains: a modulating N-terminal domain,  
CC a DNA-binding domain and a C-terminal steroid-binding domain.  
CC -|- PTM: Phosphorylated by cyclin A/CDK2 (By similarity).  
CC -|- MISCELLANEOUS: In the absence of ligand, steroid hormone receptors  
CC are thought to be weakly associated with nuclear components;  
CC hormone binding greatly increases receptor affinity. The  
CC hormone-receptor complex appears to recognize discrete DNA  
CC sequences upstream of transcriptional start sites.  
CC -|- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z37167; CAA85524.1; -.  
CC PIR; I47140; I47140.  
DR HSSP; P03372; 1HCP.  
DR GO; GO:0016585; C:chromatin remodeling complex; ISS.  
DR GO; GO:0016020; C:membrane; ISS.  
DR GO; GO:0030284; F:estrogen receptor activity; ISS.  
DR GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.  
DR GO; GO:0016049; P:cell growth; ISS.  
DR GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.  
DR GO; GO:0045839; P:negative regulation of mitosis; ISS.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.  
DR InterPro; IPR000536; Hormone rec.1lg.  
DR InterPro; IPR001292; Oestrgn\_receptor.  
DR InterPro; IPR001723; Stdhmn\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF02159; Oest\_reccep; 1.  
DR Pfam; PF00105; zF-C4; 1.  
DR PRINTS; PR00398; STRDHOMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.  
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger; Steroid-binding; Phosphorylation.  
FT DOMAIN 1 184 MODULATING.  
FT ZN\_FING 185 250 NUCLEAR RECEPTOR-TYPE.  
FT ZN\_FING 221 245 C4-TYPE.  
FT DOMAIN 251 310 C4-TYPE.  
FT DOMAIN 311 551 HINGE.  
FT DOMAIN 64 70 STEROID-BINDING.  
FT MOD\_RES 104 104 PHOSPHORYLATION (BY CDK2) (BY  
FT SIMILARITY).  
FT MOD\_RES 106 106 PHOSPHORYLATION (BY CDK2) (BY  
FT SIMILARITY).  
FT MOD\_RES 118 118 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 167 167 PHOSPHORYLATION (BY CK2) (BY SIMILARITY).  
FT MOD\_RES 537 537 PHOSPHORYLATION (BY TYR-KINASES) (BY  
FT SIMILARITY).  
SQ SEQUENCE 595 AA; 66361 MW; 0EC23DC3085BF001 CRC64;  
Query Match 29.9%; Score 713.5; DB 1; Length 595;  
Best Local Similarity 42.3%; Pred.No.6.le-45;  
Matches 159; Conservative 69; Mismatches 105; Indels 43; Gaps 10;  
QY 119 MNSMPKRLCLVCGDIASGYHYGASCEACKAFKKRTIQGNIYESCPATNECEITKRRK 178  
Db 176 MESAKETRYCAVCNDYASGYHYGWSCEGCKAFKKRSIQGHNDYMCPATNQCTIDKNRK 235  
QY 179 SCQAACRFMKCLKVGMLEKGVRLDRVRGRQ-KYKR-RIDAEN-----SPYLN 223  
Db 236 SCQAACRLRKCYEVGMKGIKRR-RGGMMLKHQRDDGEGRNEAVPBGDMRSANLWPS 294  
QY 224 POLVOPAKK-----PYNKIVSHLLVAEPEKIYAMPDPTVPDSDIKALTTLCIDLADREL 276  
Db 295 PLLIKHTKNSPVLSTLADQMSLQSAWMEIILGVVYKSLSFEDLVYADDYIMDEDQ 354  
QY 277 VTIIGWAKHIPGFSTLSLADQMSLQSAWMEIILGVVYKSLSFEDLVYADDYIMDEDQ 336  
Db 355 VHMIMWAKRVPGLDLSLHDQVHLLCAWLEILMIGLVKRSMEHPGKLLFAPNLLDRNQ 414  
QY 337 SK-LAGLLDLNNAIQLVKKYKSMKLEKEEFVTLKAIALANS-----DSMH 381  
Db 415 GKCVGEVMEIFDMLLATSSRFMRMNLQGEFFVCLKSIILLNSGVYTFLLSTLKSLEEKDH 474

```

QY      382  IEDVEAVQKLQDVLHEALQDYBAGQMEDP-RRAGKMLMTLEPLLQSTKAVQHFYNIKL 440
      |  :  |  |  |  :  |  |  |  :  |  |  |  :  |  |  |  :  |  |  |  :  |
Db      475  IHRV--LDKITDTLILHMA--KAGULTLQQQHRRLAQGLLLISHFRHMSNKGMEHLXNMKC 530

QY      441  EGKVEPMHKLFILEMLEA 456
      :  |  :  |  |  |  |  :  |  |  |  |  |  |  :  |  |  |  |  |  |  |
Db      531  KNVVEPLYDLLEMLDA 546

RESULT 10
ESRL_POEGU
ID      ESRL_POEGU      STANDARD;      PRT;      587 AA.
AC      Q91250;
DT      15-JUL-1999 (Rel. 38, Created)
DT      15-JUL-1999 (Rel. 38, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).
GN      ESRL OR NR3A1 OR ESR.
OS      Porphila guttata (Zebra finch) (Taeniopygia guttata).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Passeriformes; Estrildidae;
OC      Estrildinae; Taeniopygia.
OX      NCBI_TaxID=59729;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Hypothalamus;
RX      MEDLINE=97163508; PubMed=9010328;
RA      Jacobs E.C., Arnold A.P., Campagnoni A.T.;
RT      "Zebra finch estrogen receptor cDNA: cloning and mRNA expression.";
RL      J. Steroid Biochem. Mol. Biol. 59:135-145(1996).
CC      -!- FUNCTION: The steroid hormones and their receptors are involved in
CC      the regulation of eukaryotic gene expression and affect cellular
CC      proliferation and differentiation in target tissues.
CC      -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with ER-
CC      beta (By similarity).
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC      a DNA-binding domain and a C-terminal steroid-binding domain.
CC      -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC      subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L79911; AAB81108.1; -.
DR      HSSP; P03372; 1HCP.
DR      GO; GO:0016585; C:chromatin remodeling complex; ISS.
DR      GO; GO:0016020; C:membrane; ISS.
DR      GO; GO:0030284; F:estrogen receptor activity; ISS.
DR      GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.
DR      GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
DR      InterPro; IPR000536; Hormone_rec_lig.
DR      InterPro; IPR001292; Oestrgn_receptor.
DR      InterPro; IPR001723; Stdhmn_receptor.
DR      InterPro; IPR008946; Str_ncl_receptor.
DR      InterPro; IPR001628; Znf_C4steroid.
DR      Pfam; PF00104; hormone_rec; 1.
DR      Pfam; PF02159; Oest_recep; 1.
DR      Pfam; PF00105; zf-C4; 1.
DR      PRINTS; PR00398; STRDHOMONER.
DR      PRINTS; PR00047; STROIDFINGER.
DR      ProDom; PD000035; Znf_C4steroid; 1.
DR      SMART; SM00430; HOL1; 1.
DR      SMART; SM00399; Znf_C4; 1.
DR      PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW      Receptor; Transcription regulation; DNA-binding; Nuclear protein;

```

KM	Zinc-finger; Steroid-binding.
FT	DOMAIN 1 176 MODULATING.
FT	DNA_BIND 177 242 NUCLEAR RECEPTOR-TYPE.
FT	ZN_FING 177 197 C4-TYPE.
FT	ZN_FING 213 237 C4-TYPE.
FT	DOMAIN 243 302 HINGE.
FT	DOMAIN 303 587 STEROID-BINDING.
SQ	SEQUENCE 587 AA; 66553 MW; 2B254168A7A910AB CRC64;
 Query Match 29.8%; Score 712.5; DB 1; Length 587; Best Local Similarity 36.1%; Pred. No. 7.1e-45; Matches 171; Conservative 90; Mismatches 144; Indels 69; Gaps 14;	
QY	33 SSCSSF--IKTEPSSPASLTDS-----VNHS---FGSSDASGSYSTMGHQNGLD 80
Db	84 SSLAGFHLNSVPSPPYVLQTAPHWSPFIHHSSQQVPYYLIENDQGSF-----GMREA-- 136
QY	81 SPPLYPAPILGSGSPVKLYDDCSSTIVEDPTKCE----YMLNSMPKRLLVCGDIAS 136
Db	137 APPAF-----YRPNSDNRRHSTRERMSGANEGKSLSWESTKETRYCAVCNDIYAS 185
QY	137 GYHYGVASCFAAKFEKRTIQGNIEYSCPATNECETTKRRKSCQAACRFMKCLKYGMLKE 196
Db	186 GYHYGVWSCGECKAFKRSIQGHNDYMCPATNQCTIDKNRKSKCOACRLRKCYEIVGMKG 245
QY	197 GVRIDRVGRG-QKYKRIDAENSP-----YINPOLVOPAKK-----PY 234
Db	246 GIRKDR-RGGRAVMQRQREEQDSRNGBASSTELRAPDTLMASPLVVKHNKKNSPALSLTA 304
QY	235 NKIVSHLVAEPKIYAMPDFTVPDSDIKALTTLCUADRELVIIGMAKHIPGFSTISL 294
Db	305 EQMVSALEAEPLIYSEYDENRPFNEASMVTLLTNLADELVHMIMNAKRVPGFVDITL 364
QY	295 ADQMSLLQSAWMETILLGVYRSLSFEDLVYAADDYIMDEDQS-KLAGLLDLNNAITQLV 353
Db	365 HDQVHLLECAMELITIMIGLVWRSMHEHPGKLFPAPULLDRNQGCVEGMEIFDMILTATA 424
QY	354 KKKYSMTLEKEEFVTLKAIALAN-----SDSMHIIDEVAOVOKLDVILHEALQDYBAG 405
Db	425 ARFRMNTLOGEEFVCLKSIILLNSGVYTFLLSITLKSLEEKDYIHRVLDKITDTLTHLMAX 484
QY	406 QHM--EDPRRAGKMULTPLLRQTSKAVQHFNINKLEGKVPMHKLFLEMLEA 456
Db	485 SGLSLQQQHRRLAQLLLLISHIRHMSNKGMEHLNMKCKNVVPLYDLLLEMLDA 538
 RESULT 11 ESR1 MESAU ID ESR1 MESAU STANDARD; PRT; 595 AA. AC Q9QZJ5; Q9QZG6; DT 16-OCT-2001 (Rel. 40, Created) DT 16-OCT-2001 (Rel. 40, last sequence update) DT 28-FEB-2003 (Rel. 41, last annotation update) DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha). GN ESR1 OR NR3A1 OR ESR. OS Mesocricetus auratus (Golden hamster). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; OC Mesocricetus. OX NCBI_TaxID=10036; RN [1] RP SEQUENCE FROM N.A. RC TISSUE=Uterus; RX MEDLINE=20197937; PubMed=10731637; RA Bhat H.K., Vadgama J.V.; RT "Hamster estrogen receptor cDNA: cloning and mRNA expression."; RL J. Steroid Biochem. Mol. Biol. 72:47-53(2000). [2] RP SEQUENCE OF 98-291 FROM N.A. RA Jones J.E., Carpenter C.D., Lubbers L.S., Petersen S.L., Wade G.N.; RT "Return of lordosis after food deprivation and refeeding in Syrian hamsters."; RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.	





RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.",  
RL Nature 420:563-573(2002).  
RN [3]  
RP SEQUENCE OF 1-22 FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Liver;  
RX MEDLINE=20359284; PubMed=10899303;  
RA Kos M., O'Brien S., Flouriot G., Gannon F.,  
RT "Tissue-specific expression of multiple mRNA variants of the mouse  
RT estrogen receptor alpha gene.",  
RL FEBS Lett. 477:15-20(2000).  
RN [4]  
RP SEQUENCE OF 269-599 FROM N.A.  
RC STRAIN=SJL/J, and B10.S/J; TISSUE=Spleen;  
RA Ma R.Z., Teuscher C.,  
RT "Screening for candidate genes of mouse autoimmune diseases.",  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP CARBOHYDRATE-LINKAGE SITE SER-575.  
RX MEDLINE=97153020; PubMed=8999954;  
RA Jiang M.S., Hart G.W.,  
RT "A subpopulation of estrogen receptors are modified by O-linked  
RT N-acetylglucosamine.",  
RL J. Biol. Chem. 272:2421-2428(1997).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITES SER-10; THR-50 AND SER-575.  
RX MEDLINE=21124487; PubMed=11226831;  
RA Cheng X., Hart G.W.,  
RT "Glycosylation of the murine estrogen receptor-alpha.",  
RL J. Steroid Biochem. Mol. Biol. 75:147-158(2000).  
RN [7]  
RP INTERACTION WITH NCOA3.  
RX MEDLINE=97336097; PubMed=9192892;  
RA Torchia J., Rose D.W., Inostroza J., Kamei Y., Westin S., Glass C.K.,  
RA Rosenfeld M.G.,  
RT "The transcriptional co-activator p/CIP binds CBP and mediates  
RT nuclear-receptor function.",  
RL Nature 387:677-684(1997).  
RN [8]  
RP INTERACTION WITH NCOA6.  
RX MEDLINE=20250907; PubMed=10788465;  
RA Zhu Y.-J., Kan L., Qi C., Kanwar Y.S., Yeldandi A.V., Rao M.S.,  
RA Reddy J.K.,  
RT "Isolation and characterization of peroxisome proliferator-activated  
RT receptor (PPAR) interacting protein (PRIP) as a coactivator for  
RT PPAR.",  
RL J. Biol. Chem. 275:13510-13516(2000).  
RN [9]  
RP FUNCTION: Nuclear hormone receptor. The steroid hormones and their  
CC receptors are involved in the regulation of eukaryotic gene  
CC expression and affect cellular proliferation and differentiation  
CC in target tissues.  
CC -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with  
CC ESR2. Interacts with NCOA3 and NCOA6 coactivators, leading to a  
CC strong increase of transcription of target genes. Interacts with  
CC NCOA5 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,  
CC a DNA-binding domain and a C-terminal steroid-binding domain.  
CC -!- PTM: Phosphorylated by cyclin A/CDK2 (By similarity).  
CC -!- MISCELLANEOUS: In the absence of ligand, steroid hormone receptors

CC are thought to be weakly associated with nuclear components;  
CC hormone binding greatly increases receptor affinity. The  
CC hormone-receptor complex appears to recognize discrete DNA  
CC sequences upstream of transcriptional start sites.  
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M38651; AAA37580.1; -.  
DR EMBL; AK036627; BAC29510.1; -.  
DR EMBL; AK041525; BAC30973.1; -.  
DR EMBL; AJ276597; CAB85618.1; -.  
DR EMBL; AF128221; AAF22562.1; -.  
DR EMBL; AF128220; AAF22561.1; -.  
DR PIR; A40061; QRMSE.  
DR HSSP; P03372; IHCP.  
DR TRANSFAC; T00259; -.  
DR GlycoSuiteDB; P19785; -.  
DR MGD; MGI:1352467; Esrl.  
DR GO; GO:0016585; C:chromatin remodeling complex; ISS.  
DR GO; GO:0016020; C:membrane; ISS.  
DR GO; GO:0030284; F:estrogen receptor activity; ISS.  
DR GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.  
DR GO; GO:0016049; P:cell growth; NAS.  
DR GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.  
DR GO; GO:0045839; P:negative regulation of mitosis; NAS.  
DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; ISS.  
DR InterPro; IPR000536; Hormone\_rec\_lig.  
DR InterPro; IPR001292; Oestrgn\_receptor.  
DR InterPro; IPR001723; Stdhnm\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF02159; Oest\_recep; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.  
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger; Steroid-binding; Phosphorylation; Glycoprotein.  
FT DOMAIN 1 188  
FT DNA\_BIND 189 254  
FT ZN\_FING 189 209  
FT ZN\_FING 225 249  
FT DOMAIN 255 314  
FT DOMAIN 315 599  
FT DOMAIN 64 72  
FT MOD\_RES 108 108  
FT MOD\_RES 110 110  
FT MOD\_RES 122 122  
FT MOD\_RES 171 171  
FT MOD\_RES 541 541  
FT CARBOHYD 10 10  
FT CARBOHYD 50 50  
FT CARBOHYD 575 575  
FT VARIANT 591 591  
FT CONFLICT 269 269  
SQ SEQUENCE 599 AA; 66955 MW; 05F5E2FC21CC0A8B CRC64;

Query Match 29.7%; Score 710; DB 1; Length 599;  
Best Local Similarity 41.8%; Pred. No. 1.1e-44;  
Matches 157; Conservative 72; Mismatches 104; Indels 43; Gaps 10;  
QY 119 MLNSMPKRLCLVCGDIASGYHYGASCEACKAFPKRTIOGNIIEYSCPATNECEITKRRK 178  
DB 180 MESAKETRYCAVCNDYASGYHYGWSCEGCKAFFKRSIOGHNDYCPATNOCTIDKRRK 239  
QY 179 SCQACRFMKCLKVGMKEGVRIDVRGGRO-KYKR-RIDAEN-----SPYLN 223  
DB 240 SCQACRLRKCYEVMKKGIRKDR-RGGRMLKHKRQRDDLEGRENEMGASGDMRAANLMP 298  
QY 224 POLVQPAKK-----PYNKIVSHLLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADREL 276  
DB 299 PLVIKHTKNSPALSLTADQVMSALDAEPPMITYSEYDPSRPFSEASMMGLLTNLADREL 358  
QY 277 VVIIGWAKHIPGFSTLSLADQMSILOSAWMEIILGVVYRSLSFEDELVYADDYIMDEDQ 336  
DB 359 VHMIMNAKRVPGFGLNLHDQVHLLECAWLEILMIGLVWSMEHPKLLFAFNLLDRNQ 418  
QY 337 SK-LAGLIDLNNAIQLVKYKYSKMLEKEEFVTLKAIALANS-----DSMH 381  
DB 419 GKCVGEWEIFDMLLATSSRFMRNMLQGEFEVCLKSIIILNSGVYTFLSLTKSLEKDH 478  
QY 382 IEDVEAVOKLQDVHLHEALQDYEAGOHMEDP-RRAGKMLMTLPLLRQTSFKAVQHFYNIKL 440  
DB 479 IHRV-LDKITDTLIHLMA--KAGLTLQOQHRRLAQLLLILSHIRHMSNKGMEHLNMMKC 534  
QY 441 EGKVPMHKLFLEMLEA 456  
DB 535 KNVPLYDLLEMLDA 550

RESULT 13  
ESR1\_HORSE STANDARD; PRT; 594 AA.

ID ESR1\_HORSE  
AC Q9TV98;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).  
GN ESR1 OR NR3A1 OR ESR.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA McDowell K.J., Adams M.H., Green M.L., Cleaver B.D., Sharp D.C.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Nuclear hormone receptor. The steroid hormones and their  
CC receptors are involved in the regulation of eukaryotic gene  
CC expression and affect cellular proliferation and differentiation  
CC in target tissues (By similarity).  
CC -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with  
CC ESR2. Interacts with NCOA3, NCOA5 and NCOA6 coactivators, leading  
CC to a strong increase of transcription of target genes (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,  
CC a DNA-binding domain and a C-terminal steroid-binding domain.  
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3  
CC subfamily.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF124093; AAD17316.1; -.

DR HSSP; P03372; 1HCP.  
DR GO; GO:0016585; C:chromatin remodeling complex; ISS.  
DR GO; GO:0016020; C:membrane; ISS.  
DR GO; GO:0030284; F:estrogen receptor activity; ISS.  
DR GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.  
DR GO; GO:0016049; P:cell growth; ISS.  
DR GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.  
DR GO; GO:0045839; P:negative regulation of mitosis; ISS.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.  
DR InterPro; IPR005536; Hormone\_rec\_1ig.  
DR InterPro; IPR001292; Oestrgn\_receptor.  
DR InterPro; IPR001723; Strdrn\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF02159; Oest\_recep; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHOMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger; Steroid-binding; phosphorylation.  
FT DOMAIN 1 184 NUCLEAR RECEPTOR-TYPE.  
FT DNA\_BIND 185 250 C4-TYPE.  
FT ZN\_FING 185 205 C4-TYPE.  
FT ZN\_FING 221 245 HINGE.  
FT DOMAIN 251 310 STEROID-BINDING.  
FT DOMAIN 311 550 POLY-ALA.  
FT DOMAIN 64 70 PHOSPHORYLATION (BY CDK2) (BY  
FT MOD\_RES 104 104 SIMILARITY).  
FT MOD\_RES 106 106 PHOSPHORYLATION (BY CDK2) (BY  
FT MOD\_RES 106 106 SIMILARITY).  
FT MOD\_RES 118 118 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 167 167 PHOSPHORYLATION (BY CK2) (BY SIMILARITY).  
FT MOD\_RES 536 536 PHOSPHORYLATION (BY TYR-KINASES) (BY  
FT FT SIMILARITY).  
SQ SEQUENCE 594 AA; 66103 MW; DD36CA7C24C74B95 CRC64;

Query Match 29.5%; Score 704; DB 1; Length 594;  
Best Local Similarity 41.1%; Pred. No. 3.1e-44;  
Matches 154; Conservative 72; Mismatches 107; Indels 42; Gaps 9;

QY 119 MLNSMPKRLCLVCGDIASGYHYGASCEACKAFPKRTIOGNIIEYSCPATNECEITKRRK 178  
DB 176 MESAKETRYCAVCNDYASGYHYGWSCEGCKAFFKRSIOGHNDYCPATNOCTIDKRRK 235  
QY 179 SCQACRFMKCLKVGMKEGVRIDVRGGRO-KYKRRIID-----AENSP---YLN 223  
DB 236 SCQACRLRKCYEVMKKGIRKDR-RGGRMLKHKRQRDDLEGRENEMGASGDMRAANLMP 294  
QY 224 POLVQPAKK-----PYNKIVSHLLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADREL 276  
DB 295 PLVIKHTKNSPALSLTADQVMSALDAEPPVLYSEYDATRPFNEASMMGLLTNLADREL 354  
QY 277 VVIIGWAKHIPGFSTLSLADQMSILOSAWMEIILGVVYRSLSFEDELVYADDYIMDEDQ 336  
DB 355 VHMIMNAKRVPGFVDLSLHDQVHLLECAWLEILMIGLVWSMEHPKLLFAFNLLDRNQ 414  
QY 337 SK-LAGLIDLNNAIQLVKYKYSKMLEKEEFVTLKAIALANS-----DSMH 381  
DB 415 GKCVGEWEIFDMLLATSSRLRMNMLQGEFEVCLKSIIILNSGVYTFLSLTKSLEKDH 474  
QY 382 IEDVEAVOKLQDVHLHEALQDYEAGOHMEDPRAAGKMLMTLPLLRQTSFKAVQHFYNIKL 441  
DB 475 IHRV-LDKMTDTLIHLMA--KAGLTLQOQHRRLAQLLLILSHIRHMSNKGMEHLNMMKC 530  
QY 442 GKVPMHKLFLEMLEA 456  
DB 531 NVVPLYDLLEMLDA 545

RESULT 14  
 ESRI HUMAN STANDARD; PRT; 595 AA.  
 ID ESRI\_HUMAN  
 AC P03372; Q13511; Q14276; Q9NU51; Q9UDZ7; Q9UIS7;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).  
 GN ESRI OR NR3A1 OR ESR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RX MEDLINE=86122927; PubMed=3753802;  
 RA Greene G.L., Gilna P., Waterfield M., Baker A., Hort Y., Shine J.;  
 RT "Sequence and expression of human estrogen receptor complementary  
 RT DNA.";  
 RL Science 231:1150-1154(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RX MEDLINE=86146892; PubMed=3754034;  
 RA Green S., Walter P., Kumar V., Krust A., Bornert J.-M., Argos P.,  
 RT Chambon P.;  
 RT "Human oestrogen receptor cDNA: sequence, expression and homology to  
 RT v-erb-A.";  
 RL Nature 320:134-139(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG), AND VARIANT ASP-411 INS.  
 RC TISSUE=Breast;  
 RX MEDLINE=96174665; PubMed=8600466;  
 RA Pink J.J., Wu S.Q., Wolf D.M., Bilimoria M.M., Jordan V.C.;  
 RT "A novel 80 kDa human estrogen receptor containing a duplication of  
 RT exons 6 and 7.";  
 RL Nucleic Acids Res. 24:962-969(1996).  
 RN [4]  
 RP SEQUENCE OF 152-595 FROM N.A. (ISOFORM LONG).  
 RX MEDLINE=20084372; PubMed=10619354;  
 RA Schubert E.L., Lee M.K., Newman B., King M.C.;  
 RT "Single nucleotide polymorphisms (SNPs) in the estrogen receptor gene  
 RT and breast cancer susceptibility.";  
 RL J. Steroid Biochem. Mol. Biol. 71:21-27(1999).  
 RN [5]  
 RP SEQUENCE OF 216-434 FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Breast carcinoma;  
 RX MEDLINE=93153765; PubMed=7916651;  
 RA Pfeiffer U., Recarotta E., Castagnetta L., Vidali G.;  
 RT "Estrogen receptor variant messenger RNA lacking exon 4 in estrogen-  
 RT responsive human breast cancer cell lines.";  
 RL Cancer Res. 53:741-743(1993).  
 RN [6]  
 RP SEQUENCE OF 110-117, PHOSPHORYLATION, AND MUTAGENESIS.  
 RX MEDLINE=96026869; PubMed=7476978;  
 RA Joel P.B., Traish A.M., Lannigan D.A.;  
 RT "Estradiol and phorbol ester cause phosphorylation of serine 118 in  
 RT the human estrogen receptor.";  
 RL Mol. Endocrinol. 9:1041-1052(1995).  
 RN [7]  
 RP SEQUENCE OF 354-548 FROM N.A.  
 RC TISSUE=Breast carcinoma;  
 RA Naundorf H., Becker M., Fiebig C., Buettner B., Fichtner I.;  
 RT "Mechanisms of acquired tamoxifen resistance in a xenotransplanted  
 RT human breast carcinoma.";  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE OF 413-595 FROM N.A.  
 RA Parker A.;  
 RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE OF 532-542, AND PHOSPHORYLATION.

RX MEDLINE=95280953; PubMed=7539106;  
 RA Arnold S.F., Obourn J.D., Jaffe H., Notides A.C.;  
 RT "Phosphorylation of the human estrogen receptor on tyrosine 537 in  
 RT vivo and by src family tyrosine kinases in vitro.";  
 RL Mol. Endocrinol. 9:24-33(1995).  
 RN [10]  
 RP INTERACTION WITH NCOA6.  
 RX MEDLINE=20036574; PubMed=10567404;  
 RA Lee S.-K., Anzick S.L., Choi J.-E., Bubendorf L., Guan X.-Y.,  
 RA Jung Y.-K., Kallioniemi O.P., Kononen J., Trent J.M., Azorsa D.,  
 RA Jhun B.-H., Cheong J.H., Lee Y.C., Meltzer P.S., Lee J.W.;  
 RT "A nuclear factor ASC-2, as a cancer-amplified transcriptional  
 RT coactivator essential for ligand-dependent transactivation by nuclear  
 RT receptors in vivo.";  
 RL J. Biol. Chem. 274:34283-34293(1999).  
 RN [11]  
 RP INTERACTION WITH NCOA5.  
 RX MEDLINE=20565767; PubMed=11113208;  
 RA Sauve F., McBroome L.D.B., Gallant J., Moraitis A.N., Labrie F.,  
 RA Giguere V.;  
 RT "CIA, a novel estrogen receptor coactivator with a bifunctional  
 RT nuclear receptor interacting determinant.";  
 RL Mol. Cell. Biol. 21:343-353(2001).  
 RN [12]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=95140025; PubMed=7838153;  
 RA Arnold S.F., Obourn J.D., Jaffe H., Notides A.C.;  
 RT "Serine 167 is the major estradiol-induced phosphorylation site on the  
 RT human estrogen receptor.";  
 RL Mol. Endocrinol. 8:1208-1214(1994).  
 RN [13]  
 RP PHOSPHORYLATION OF SER-104 AND SER-106, AND MUTAGENESIS.  
 RX MEDLINE=99357754; PubMed=10428798;  
 RA Rogatsky I., Trowbridge J.M., Garabedian M.J.;  
 RT "Potential of human estrogen receptor alpha transcriptional  
 RT activation through phosphorylation of serines 104 and 106 by the  
 RT cyclin A-CDK2 complex.";  
 RL J. Biol. Chem. 274:22296-22302(1999).  
 RN [14]  
 RP STRUCTURE BY NMR OF 180-262.  
 RX MEDLINE=91061891; PubMed=2247153;  
 RA Schwabe J.W.E., Neuhaus D., Rhodes D.;  
 RT "Solution structure of the DNA-binding domain of the oestrogen  
 RT receptor.";  
 RL Nature 348:458-461(1990).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 180-262.  
 RX MEDLINE=94037103; PubMed=8221895;  
 RA Schwabe J.W.E., Chapman L., Finch J.T., Rhodes D.;  
 RT "The crystal structure of the estrogen receptor DNA-binding domain  
 RT bound to DNA: how receptors discriminate between their response  
 RT elements.";  
 RL Cell 75:567-578(1993).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 305-548.  
 RX MEDLINE=97478539; PubMed=9338790;  
 RA Brzozowski A.M., Pike A.C.W., Dauter Z., Hubbard R.E., Bonn T.,  
 RA Engstrom O., Oehman L., Greene G.L., Gustafsson J.-A., Carlquist M.;  
 RT "Molecular basis of agonism and antagonism in the oestrogen  
 RT receptor.";  
 RL Nature 389:753-758(1997).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 306-544.  
 RX MEDLINE=98263297; PubMed=9600906;  
 RA Tanenbaum D.M., Wang Y., Williams S.P., Sigler P.B.;  
 RT "Crystallographic comparison of the estrogen and progesterone  
 RT receptor's ligand binding domains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:5998-6003(1998).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 294-554.  
 RX MEDLINE=99091051; PubMed=9875847;  
 RA Shiau A.K., Barstad D., Loria P.M., Cheng L., Kushner P.J.,  
 RA Agard D.A., Greene G.L.;





DR InterPro; IPR008946; Str ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF02159; Oest\_recep; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOLI; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger; Steroid-binding.  
FT DOMAIN 1 179 MODULATING.  
FT DNA\_BIND 180 245 NUCLEAR RECEPTOR-TYPE.  
FT ZN\_FING 180 200 C4-TYPE.  
FT ZN\_FING 216 240 C4-TYPE.  
FT DOMAIN 246 302 HINGE.  
FT DOMAIN 303 586 STEROID-BINDING.  
SQ SEQUENCE 586 AA; 66080 MW; 0EDC77EB0D6F08BF CRC64;

Query Match 29.1%; Score 694; DB 1; Length 586;  
Best Local Similarity 34.5%; Pred. No. 1.6e-43;  
Matches 161; Conservative 86; Mismatches 140; Indels 80; Gaps 12;

QY 43 PSSPA-----SLTDSVNHHS---PGSSSDASGSY-----SSTMNGHONGLD 80  
Db 99 PPSPVVFLAKLPQLSPFIHHGQOVYYLSEQGTFAVREAPPTFYRSSSDNRROSGRE 158  
QY 81 SPPLYPAPILIGSGSPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHY 140  
Db 159 R-----MSSANDKGPPS---WESTKETRYCAVCSDYASGYHY 192  
QY 141 GVASCEACKAFKRTIOGNIIEYSCPATNECEITKRRKSCQACRFMKCLKYGMKKEGVRL 200  
Db 193 GWSCEGCKAFKRTIOGNDYMCAPTNOCTIDKNRRKSCQACRLRKYEVGMKKGIRK 252  
QY 201 DRVRRGR-----QYKRRIDAEN---SPYINPQLVQPAKKPY-----NKIVSHL 241  
Db 253 DR-RGRLKHKRQKEEQEQKNDVDPSFIRTAStWNPVSVMKLSPLVSLTAEQLLISAL 311  
QY 242 LVAEPEKIYAMPDPTVPDSIDIKALTTLCDIADRELVTIIGWAKHIPGFSTISLADQMSL 301  
Db 312 MEAEAPIVYSEHDSSTKFLSEASMTLLTNLADRELVMINMAKRVPGFVDLTIHDQVHL 371  
QY 302 QSAMEITLIGVYRSLSEFEDLVYADDYIMEDQSK-LAGLLDINNAILQLVKYKYSMK 360  
Db 372 ECAYLEILMVGLWRSVEHPKLSFAPNLLIDRNQRCVEGLVEIFDMLVTATFRMR 431  
QY 361 IEKEEFVTLKATALANSDSM-----HIEDVEAVQKLQDVLHEALQDYEAQGM---E 409  
Db 432 LRGEETFCLKSIIILNSGVYTFLLSSTLESLEDTLIHILDKIIDTLVHFMAKSGLSIQ 491  
QY 410 DPRRAGKMLTLLRLRQSTKAVQHFYNIKLEGVPMHKLFLFMLFA 456  
Db 492 QQRRLAQLLILLSHIRMSNKGMEHLYSMCKKNVPLIYDLLLEMLDA 538

Search completed: August 3, 2004, 12:40:55  
Job time : 14 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:38:06 ; Search time 41 Seconds

(without alignments)  
3524.569 Million cell updates/sec

Title: US-10-054-841-4

Perfect score: 2388

Sequence: 1 MDSVELCLPESPSLHYEEL.....KLEGVPMFKLFLEMLBAKV 458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2282	95.6	466	11 Q8CHC9	Q8chc9 mus musculu
2	1756	73.5	434	11 Q80VSI	Q80vsi1 mus musculu
3	1751	73.3	433	11 Q8C7A6	Q8c7a6 mus musculu
4	1744	73.0	433	11 Q8CCV5	Q8ccv5 mus musculu
5	1106	46.3	262	4 Q9HCB2	Q9hcb2 homo sapien
6	1054.5	44.2	323	4 Q96102	Q96102 homo sapien
7	936.5	39.2	299	4 Q8N4S8	Q8n4s8 homo sapien
8	739.5	31.0	484	5 Q9VSE9	Q9vse9 drosophila
9	739.5	31.0	496	5 Q8WS79	Q8ws79 drosophila
10	726	30.4	587	13 Q8UWB0	Q8uwb0 caiman croc
11	724	30.3	589	13 Q8AYH0	Q8ayh0 coturnix co
12	715	29.9	581	13 Q8UWA9	Q8uwa9 cnemidophor
13	697.5	29.2	152	13 Q8QG80	Q8qg80 brachydanio
14	695	29.1	574	13 Q7T2K8	Q7t2k8 halichoeres
15	691	28.9	129	4 Q9UNJ4	Q9unj4 homo sapien
16	690	28.9	554	13 Q90ZM8	Q90zm8 petromyzon

17	689	28.9	620	13 Q90WH6	Q90wh6 clarias gar
18	684	28.6	583	13 Q804Q6	Q804q6 acanthopagr
19	683.5	28.6	542	13 Q90ZE6	Q90ze6 squalus aca
20	682.5	28.6	431	6 Q95L13	Q95l13
21	681	28.5	564	13 Q90WV1	Q90wv1 carassius a
22	678	28.4	620	13 Q7S210	Q7sz10 fundulus he
23	677	28.4	578	13 Q8QHL0	Q8qhl0 paralichthy
24	675.5	28.3	570	13 Q800Q2	Q800q2 zoarces viv
25	664.5	27.8	553	13 Q7ZUJ2	Q7zu32 brachydanio
26	663.5	27.8	553	13 Q98SM7	Q98sm7 brachydanio
27	663.5	27.8	553	13 Q90MS8	Q90ws8 brachydanio
28	659.5	27.6	559	13 Q8UTB9	Q8jtb9 cyprinus ca
29	653.5	27.4	553	13 Q8AV62	Q8ave2 brachydanio
30	648.5	27.2	671	13 Q7T2K7	Q7t2k7 halichoeres
31	646.5	27.1	458	13 Q9DDJ3	Q9ddj3 halichoeres
32	645	27.0	553	13 Q804Q7	Q804q7 acanthopagr
33	637	26.7	562	13 Q8UW75	Q8uw75 oryzias lat
34	636	26.6	567	11 Q8BG65	Q8bg65 mus musculu
35	632	26.5	503	11 Q91Z86	Q91z86 mus musculu
36	629	26.3	592	13 Q90WS9	Q90ws9 brachydanio
37	624.5	26.2	565	13 Q8QHK9	Q8qhk9 paralichthy
38	624.5	26.2	612	13 Q7T3U4	Q7t3u4 varicorhinu
39	621	26.0	592	13 Q98SM8	Q98sm8 brachydanio
40	610.5	25.6	601	13 Q7T3U5	Q7t3u5 candidia ba
41	600.5	25.1	436	5 Q8T5C6	Q8t5c6 biophthalmi
42	595	24.9	522	5 Q8MX78	Q8mx78 brachiosto
43	588.5	24.6	363	5 Q9UAF1	Q9uaf1 polyanthroca
44	586.5	24.6	499	6 Q95MF0	Q95mf0 macaca arc
45	584	24.5	412	13 Q91840	Q91840 xenopus lae

ALIGNMENTS

RESULT 1

ID	Q8CHC9	PRELIMINARY;	PRT;	466 AA.
AC	Q8CHC9;			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	MKIAA0832 protein (Fragment).			
GN	ESRRG OR MKIAA0832.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	Okazaki N., Kikuno R., Ohara R., Inamoto S., Hara Y., Nagase T.,			
RA	Ohara O., Koga H.;			
RT	"Prediction of the coding sequences of mouse homologues of KIAA gene:			
RT	I. The complete nucleotide sequences of 100 mouse KIAA-homologous			
RT	CDNAs identified by screening of terminal sequences of cDNA clones			
RT	randomly sampled from size-fractionated libraries.";			
RL	Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; AB093266; BAC41450.1; -.			
DR	MGD; MGI:1347056; Esrrg.			
DR	GO; GO:0005634; C:nucleus; IEA.			
DR	GO; GO:0003707; F:steroid hormone receptor activity; IEA.			
DR	GO; GO:0003700; F:transcription factor activity; IEA.			
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.			
DR	InterPro; IPR000536; Hormone_rec_lig.			
DR	InterPro; IPR001723; Stdrmn_receptor.			
DR	InterPro; IPR008946; Str_ncl_receptor.			
DR	InterPro; IPR001628; Znf_C4steroid.			
DR	Pfam; PF00104; hormone_rec_1.			
DR	Pfam; PF00105; zf-C4; 1.			
DR	PRINTS; PR00398; STRDHORMONER.			
DR	PRINTS; PR00047; STROIDFTNGER.			
DR	ProDom; PD000035; Znf_C4steroid; 1.			
DR	SMART; SM00430; HOL1; 1.			

DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
FT NON\_TER 1  
SQ SEQUENCE 466 AA; 52006 MW; 1EB8C98B180D6FEF CRC64;

Query Match 95.6%; Score 2282; DB 11; Length 466;  
Best Local Similarity 99.8%; Pred. No. 1e-181;  
Matches 438; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 LLCRMSNDRHIDSSCSSFIKTEPSSASLTDVNVHSPGSSSDASGYSSTMNGHONGL 79  
DB 28 LLCRMSNDRHIDSSCSSFIKTEPSSASLTDVNVHSPGSSSDASGYSSTMNGHONGL 87  
QY 80 DSPPLYPAPILGSGPVRKLYDDCSSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYH 139  
DB 88 DSPPLYPAPILGSGPVRKLYDDCSSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYH 147  
QY 140 YGVASCEACKAFKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMLKEGVR 199  
DB 148 YGVASCEACKAFKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMLKEGVR 207  
QY 200 LDRVGRGRQKRRIDAENSPYLNQIVQPAKRPYNKIVSHLLVAEPEKIYAMPDPTVPD 259  
DB 208 LDRVGRGRQKRRIDAENSPYLNQIVQPAKRPYNKIVSHLLVAEPEKIYAMPDPTVPD 267  
QY 260 SDIKALTTLCDLADRELVIIGWAKHIFGFSTLSLADQMSLLQSAWMEILLIGVYRSLS 319  
DB 268 SDIKALTTLCDLADRELVIIGWAKHIFGFSTLSLADQMSLLQSAWMEILLIGVYRSLS 327  
QY 320 FEDELVYADDYIMDEDQSLAGLLDNNAILQLVKKYSMKLEKEEFVTLKALALANSDS 379  
DB 328 FEDELVYADDYIMDEDQSLAGLLDNNAILQLVKKYSMKLEKEEFVTLKALALANSDS 387  
QY 380 MHIEDVEAVOKLQDVLHEALQDYEAQGMEDPRRAGKMLMTLLPLLRQSTKAVQHFYNIK 439  
DB 388 MHIEDVEAVOKLQDVLHEALQDYEAQGMEDPRRAGKMLMTLLPLLRQSTKAVQHFYNIK 447  
QY 440 LEGKVPMHKLFLEMLEAKV 458  
DB 448 LEGKVPMHKLFLEMLEAKV 466

RESULT 2

Q80VS1 PRELIMINARY; PRT; 434 AA.

ID Q80VS1  
AC Q80VS1;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Similar to estrogen related receptor, beta (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Eye;  
RA Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC044858; AAH44858.1; -.  
DR PIR; PT0706; PT0706.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR00536; Hormone\_rec\_lig.  
DR InterPro; IPR001723; Strhma\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR01628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDEFINGER.

DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOLI; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
FT NON\_TER 1  
SQ SEQUENCE 434 AA; 48484 MW; 3679E7337E5762DC CRC64;

Query Match 73.5%; Score 1756; DB 11; Length 434;  
Best Local Similarity 75.0%; Pred. No. 7e-138;  
Matches 327; Conservative 55; Mismatches 52; Indels 2; Gaps 1;

QY 23 RMSNKRHIDSSCSSFIKTEPSSASLTDVNVHSPGSSSDASGYSSTMNGHONGL 82  
DB 1 RMSEDRHLGSSCGSFIKTEPSSSGIDALSHSPSGSSDASGFGIALSTHANGLDSP 60  
QY 83 PLYPAPILGSGPVRKLYDDCSSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYG 142  
DB 61 PMFAGAGL--GNPCKRSYEDCTSGIMEDSAIKCEYMLNAIPKRLCLVCGDIASGYHYG 118  
QY 143 ASCEACKAFKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMLKEGRLDR 202  
DB 119 ASCEACKAFKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMLKEGRLDR 178  
QY 203 VRGGRQKRRIDAENSPYLNQIVQPAKRPYNKIVSHLLVAEPEKIYAMPDPTVPDSI 262  
DB 179 VRGGRQKRRIDAENSPYLNQIVQPAKRPYNKIVSHLLVAEPEKIYAMPDPTVPDSI 238  
QY 263 KALTTLCDLADRELVIIGWAKHIFGFSTLSLADQMSLLQSAWMEILLIGVYRSLSFED 322  
DB 239 KALTTLCDLADRELVIIGWAKHIFGFSTLSLADQMSLLQSAWMEILLIGVYRSLSFED 298  
QY 323 ELVYADDYIMDEDQSLAGLLDNNAILQLVKKYSMKLEKEEFVTLKALALANSDSMHI 382  
DB 299 KLAVYADDYIMDEHSRYGLLDLYRAILQLVRRYKLVKEEFMTLKALALANSDSMYI 358  
QY 383 EDVEAVOKLQDVLHEALQDYEAQGMEDPRRAGKMLMTLLPLLRQSTKAVQHFYNIK 442  
DB 359 ENLEAVOKLQDVLHEALQDYELSRHEPRRAGKMLMTLLPLLRQSTKAVQHFYNIK 418  
QY 443 KVPMHKLFLEMLEAKV 458  
DB 419 KVPMHKLFLEMLEAKV 434

RESULT 3

Q8C7A6 PRELIMINARY; PRT; 433 AA.

ID Q8C7A6  
AC Q8C7A6;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Estrogen related receptor.  
GN ESRB.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Heart;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573 (2002).  
DR EMBL; AK052256; BAC34898.1; -.  
DR PIR; PT0706; PT0706.  
DR MGD; MGI:1346832; ESRrb.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR000536; Hormone\_rec\_1lg.  
DR InterPro; IPR001723; Stdhrm\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOLI; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
SQ SEQUENCE 433 AA; 48328 MW; 1A0383E7E2ED8BF3 CRC64;

Query Match 73.3%; Score 1751; DB 11; Length 433;  
Best Local Similarity 74.9%; Pred. No. 1.8e-137;  
Matches 326; Conservative 55; Mismatches 52; Indels 2; Gaps 1;

QY 24 MSNKRHIDSSCSSFIKTEPSSPASLTDVSNHSPGSSSDASGYSSTMNGHQLDSDP 83  
DB 1 MSSEDRHLGSSCGSFIKTEPSSPSGIDALSHSPSSGSDASGGFGIALSTHANGLDSDP 60  
QY 84 LYPAPILGSGPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143  
DB 61 MFAGAGL--GGNPCRSYEDCTSGIMEDSAIKCEYMLNAIPKRLCLVCGDIASGYHYGVA 118  
QY 144 SCEACKAFKKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMKEGVRLDRV 203  
DB 119 SCEACKAFKKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMKEGVRLDRV 178  
QY 204 RGGROKKRRIDAENSPYLNQVPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK 263  
DB 179 RGGROKKRRIDSENSPYLNLPISPAKKPLTKIVSNLLGVEQDKLYAMPNDIPEGDIK 238  
QY 264 ALTLCDLADRELVIIGWAKHIFGFTSLADQMSLLQSAWMEILLGVYRSLSFED 323  
DB 239 ALTLCELADRELVLINWAKHIPFPSLTLDQMSLLQSAWMEILLGVYRSLPYDK 298  
QY 324 LYYADYIMDEDQSKLAGLLDNNAILQLVKYKSMKLEKEEFVTLKALALANSDSMITE 383  
DB 299 LAYAEDYIMDEHSRLVGLLDYRALILQVRYKCLKVEKEEFMILLKALALANSDSMITE 358  
QY 384 DVEAVQKLQDLVHEALQDYEAGQHMEDPRRAGKMLMTPLRLQSTKAVQHFYNIKLEK 443  
DB 359 NLEAVQKLQDLVHEALQDYELSQHHEPRRAGKLLTLPLRLQTAAKAVQHFYSVKLQK 418  
QY 444 VPMKLFLEMLEAKV 458  
DB 419 VPMKLFLEMLEAKV 433

RESULT 4  
Q8CCV5 PRELIMINARY; PRT; 433 AA.  
ID Q8CCV5  
AC Q8CCV5;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Estrogen related receptor.  
GN ESRRB.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium;  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).  
DR EMBL; AK032025; BAC27656.1; -.  
DR PIR; PT0706; PT0706.  
DR MGD; MGI:1346832; Esrrb.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR000536; Hormone\_rec\_1lg.  
DR InterPro; IPR001723; Stdhrm\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOLI; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
SQ SEQUENCE 433 AA; 48255 MW; C802E08680ED89F2 CRC64;

Query Match 73.0%; Score 1744; DB 11; Length 433;  
Best Local Similarity 74.7%; Pred. No. 7e-137;  
Matches 325; Conservative 55; Mismatches 53; Indels 2; Gaps 1;

QY 24 MSNKRHIDSSCSSFIKTEPSSPASLTDVSNHSPGSSSDASGYSSTMNGHQLDSDP 83  
DB 1 MSSEDRHLGSSCGSFIKTEPSSPSGIDALSHSPSSGSDASGGFGIALSTHANGLDSDP 60  
QY 84 LYPAPILGSGPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143  
DB 61 MFAGAGL--GGNPCRSYEDCTSGIMEDSAIKCEYMLNAIPKRLCLVCGDIASGYHYGVA 118  
QY 144 SCEACKAFKKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMKEGVRLDRV 203  
DB 119 SCEACKAFKKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMKEGVRLDRV 178  
QY 204 RGGROKKRRIDAENSPYLNQVPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK 263  
DB 179 RGGROKKRRIDSENSPYLNLPISPAKKPLTKIVSNLLGVEQDKLYAMPNDIPEGDIK 238  
QY 264 ALTLCDLADRELVIIGWAKHIFGFTSLADQMSLLQSAWMEILLGVYRSLSFED 323  
DB 239 ALTLCELADRELVLINWAKHIPFPSLTLDQMSLLQSAWMEILLGVYRSLPYDK 298  
QY 324 LYYADYIMDEDQSKLAGLLDNNAILQLVKYKSMKLEKEEFVTLKALALANSDSMITE 383  
DB 299 LAYAEDYIMDEHSRLVGLLDYRALILQVRYKCLKVEKEEFMILLKALALANSDSMITE 358  
QY 384 DVEAVQKLQDLVHEALQDYEAGQHMEDPRRAGKMLMTPLRLQSTKAVQHFYNIKLEK 443  
DB 359 NLEAVQKLQDLVHEALQDYELSQHHEPRRAGKLLTLPLRLQTAAKAVQHFYSVKLQK 418  
QY 444 VPMKLFLEMLEAKV 458  
DB 419 VPMKLFLEMLEAKV 433

RESULT 5  
Q9HCB2 PRELIMINARY; PRT; 262 AA.  
ID Q9HCB2  
AC Q9HCB2;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE HERB2 (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.



RA Rowen L., Madan A., Qin S., Baradarani L., Birditt B., Bloom S.,  
RA Dors M., Dickhoff R., Fleetwood P., Harrison G., Kaur A., Madan A.,  
RA Nesbitt R., Traicoff R., Hood L.,  
RT "Sequencing of human chromosome 14q24.3 region."  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
DR EMBL; AC016543; AAG29619.1; -.  
DR HSSP; P03372; IHCQ.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR InterPro; IPR001723; Stchrn\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR PRODOM; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;  
KW Transcription regulation; Zinc; Zinc-finger.  
FT NON\_TER 262 262  
SQ SEQUENCE 262 AA; 28203 MW; 94FA8979AC2D43D3 CRC64;

Query Match 46.3%; Score 1106; DB 4; Length 262;  
Best Local Similarity 78.0%; Pred. No. 5.7e-84;  
Matches 206; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

QY 24 MSNKRHIDSSCSSEFIKTEPSSPATLDSVNHSPGSSDASGYSSTMNGHQLDSP 83  
DB 1 MSSDRHLGSSCGSFIKTEPSSSGIDALSHSPSGSSDASGFGALGTHANGLDSP 60  
QY 84 LYPSPALIGSGSPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCYCGDIASGYHYGA 143  
DB 61 MFAGAG-IGGT-PCRKSYEDCAGIMEDSAIKCEYMLNAIPKRLCYCGDIASGYHYGA 118  
QY 144 SCEACKAFPKRTIQGNI EYSCPATNECETITRRRKSQAACRFMKCLKVGMKEGVRLD 203  
DB 119 SCEACKAFPKRTIQGNI EYSCPATNECETITRRRKSQAACRFMKCLKVGMKEGVRLD 178  
QY 204 RGGROKYKRRIDAENSPYLPQVPAKPYNKIVSHLLVAEPEKIYAMPDPTVPDS 263  
DB 179 RGGROKYKRRIDSESSPYLSLQISSPAKPLTKIVSYLLVAEPDKLYAMPDPGMEGD 238  
QY 264 ALTTLCDLADRELVIIGWAKHIP 287  
DB 239 ALTTLCDLADRELVIIGWAKHIP 262

RESULT 6  
Q96102 PRELIMINARY; PRT; 323 AA.  
AC Q96102; 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)  
DE Hypothetical protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Uterus;  
RA Strausberg R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
DR EMBL; BC007915; AAH07915.1; -.

DR HSSP; P19793; 1LBD.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000515; BPD\_transp.  
DR InterPro; IPR000536; Hormone\_rec\_lig.  
DR InterPro; IPR001723; Stchrn\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRODOM; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOLI; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00402; BPD\_TRANS INN\_MEMBER; 1.  
KW Hypothetical protein; DNA-binding; Metal-binding; Nuclear protein;  
KW Receptor; Transcription; Transcription regulation; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
SQ SEQUENCE 323 AA; 35341 MW; 75D5BA85958025DC CRC64;

Query Match 44.2%; Score 1054.5; DB 4; Length 323;  
Best Local Similarity 63.7%; Pred. No. 1.5e-79;  
Matches 202; Conservative 50; Mismatches 52; Indels 13; Gaps 2;

QY 153 KRTIQGNI EYSCPATNECETITRRRKSQAACRFMKCLKVGMKEGVRLDVRGGRQYKR 212  
DB 4 KRTIQGSI EYSCPATNECETITRRRKAQAACRFMKCLKVGMKEGVRLDVRGGRQYKR 63  
QY 213 RIDAENSPYLPQVOP-----AKPEYNKIVSHLLVAEPEKIYAMPDPTVPDS 263  
DB 64 RPEVDPLDPFPPGPPAGPLAVAGGPKRTAAPVNALVSHLLVAEPEKIYAMPDPAGD 123  
QY 264 ALTTLCDLADRELVIIGWAKHIPGFTSLADQMSLQSAAMEILIGVYRSLSFEDE 323  
DB 124 AVATLCLDFREIVITISWAKSIPGSSLSLQSDMSVLQSVWMEVLVYGAQKSLPLQDE 183  
QY 324 LVYADYIMDEDSKLAGLLDNLNAILQVKKYSKMLEKEEFVTLKAILANSDSMHIE 383  
DB 184 LAFABDLVDEEGARAAGLGLGALLQVRLQALRLEREYVLLKAILANSDSVHIE 243  
QY 384 DVEAVQKLDVYLHEALQDYEAQ-----HMEPPRAQKMLMTPLRLQSTKAVQHFY 439  
DB 244 DAEAVEQLREALHEALLEYEAAGRAGPGGAERRRAGRLLTLPLRLQSTKAVLAHFY 303  
QY 440 LEKVPMHKLFLMLEA 456  
DB 304 LEKVPMHKLFLMLEA 320

RESULT 7  
Q8N4S8 PRELIMINARY; PRT; 299 AA.  
AC Q8N4S8; 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)  
DE Similar to estrogen-related receptor alpha (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Cervix;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
 DR EMBL; BC033701; AAH33701.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:steroid hormone receptor activity; IEA.  
 DR GO; GO:0003700; F:steroid hormone receptor activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000515; BPD\_transp.  
 DR InterPro; IPR000536; Hormone\_rec\_lig.  
 DR InterPro; IPR001723; Stdhmr\_receptor.  
 DR InterPro; IPR008946; Str\_ncl\_receptor.  
 DR InterPro; IPR001628; Znf\_C4steroid.  
 DR Pfam; PF00104; hormone\_rec; 1.  
 DR Pfam; PF00105; zf-C4; 1.  
 DR PRINTS; PR00398; STRDHORMONER.  
 DR SMART; SM00430; HOL1; 1.  
 DR PROSITE; PS00402; BPD\_TRANSF\_INN\_MEMBER; 1.  
 KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;  
 KW Transcription regulation; Zinc; Zinc-finger.  
 FT NON TER 1  
 SQ SEQUENCE 299 AA; 32629 MW; 186AC8A78B149AD9 CRC64;  
 Query Match 39.2%; Score 936.5; DB 4; Length 299;  
 Best Local Similarity 61.4%; Pred. No. 9.5e-70;  
 Matches 181; Conservative 48; Mismatches 53; Indels 13; Gaps 2;  
 QY 175 RRRKSCQACRFMKCLKVGMLKEGVRLDRVGRGQYKRRIDAENSPLYNPQLVQP----- 229  
 Db 2 RPRKACQACRFMKCLKVGMLKEGVRLDRVGRGQYKRRIDAENSPLYNPQLVQP----- 61  
 QY 230 ----AKKPYNKIVSHLVAEPKIVAMPDPVPSDILKALTTLCDLADRELVTIIGWAKH 285  
 Db 62 GPRKTAAPVNAVLVSHLLVVEPEKLYAMPDPAGDGLPAVATLCDLFPREIVVTISWAKS 121  
 QY 286 IPGFSTLSADQMSLLQSAWMEILLGVVYRSLSFEDLVYADDYIMDEQSLAGLLDL 345  
 Db 122 IPGFSTLSADQMSLVQSWMEVVLGVAGRSPLQDELAFADLVLDDEGARAAGLGEL 181  
 QY 346 NNAIIQLVKYKSMKLEKEEFTVTKAIALANSDSMHIEDVEAVQKLQDYLHEALQDYEAG 405  
 Db 182 GAALLQLVRLQLRLEREYVLLKALALANSDSVHIEDAEAVEQLREALHEALLLEYEAG 241  
 QY 406 Q----HMEPPRAGKXMTLPLLRQTSKAVQHFYNIKLEGVPMHKLPLEMLEA 456  
 Db 242 RAGPGGAERRRAGRLITLPLLRQTAGKVLAFHYGVKLEGVPMHKLPLEMLEA 296  
 RESULT 8  
 Q9VSE9 PRELIMINARY; PRT; 484 AA.  
 AC Q9VSE9; 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE CG7404 protein (GH28308P) (Estrogen-related receptor).  
 GN ERR OR CG7404.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Li X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celnik S.;  
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Lehman J., Wherry S., Talmadge C., Davis J.R., Kiss I., Eudy J.D.,  
 RA Chase B.A., Sumegi J.;  
 RT "Identification and characterization of the Drosophila ortholog of  
 RT vertebrate estrogen-related receptor gene and its product.";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
 DR EMBL; AE003556; AAP50473.1; -.  
 DR EMBL; AY051632; AAK93056.1; -.  
 DR EMBL; AF359420; AAI37553.1; -.  
 DR HSSP; P03372; 1HCQ.  
 DR FlyBase; FBgn0035849; ERR.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR InterPro; IPR000536; Hormone\_rec\_lig.  
 DR InterPro; IPR001723; Stdhmr\_receptor.  
 DR InterPro; IPR008946; Str\_ncl\_receptor.  
 DR InterPro; IPR001628; Znf\_C4steroid.  
 DR Pfam; PF00104; hormone\_rec; 1.  
 DR Pfam; PF00105; zf-C4; 1.  
 DR PRINTS; PR00398; STRDHORMONER.  
 DR PRINTS; PR00047; STROIDFINGER.  
 DR ProDom; PD000035; Znf\_C4steroid; 1.  
 DR SMART; SM00430; HOL1; 1.  
 DR SMART; SM00399; Znf\_C4; 1.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.





DR EMBL; AE003556; AAN12015.1; -  
DR FlyBase; FBgn0035849; ERR.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR InterPro; IPR000536; Hormone\_rec\_lig.  
DR InterPro; IPR001723; Steroid\_receptor.  
DR InterPro; IPR008946; Steroid\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDEFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;  
KW Transcription regulation; Zinc; Zinc-finger.  
SQ SEQUENCE 496 AA; 54529 MW; 5D67E5AE8E9DCB32 CRC64;

Query Match 31.0%; Score 739.5; DB 5; Length 496;  
Best Local Similarity 37.4%; Pred. No. 5.4e-53;  
Matches 189; Conservative 70; Mismatches 151; Indels 95; Gaps 13;

QY 14 LHYEEL-----LCRMGNKDRHIDSSCSSTFIKTEP-----SSPASLTDSVNHSPG 59  
DB 20 LHIQEVDPSPASCFSPSSKSTATQSGTNGLKSSPSVSPEROLCSTSTSLCDLHNVSLS 79  
QY 60 GSSDA-SGSYSSTMN-GHONGLDSPPLPSAPILGSGCPVRKLYDDCSSTIVEDPQTKCE 117  
DB 80 NDGDSLKSGSTGGNGGGGGGTSGGNATNAGAGSGSVR--DEL----- 123  
QY 118 YMLNSMPKRLCYCGDIASGYHYGASCEACKAFKRTIOGNIIEYSCPATNECEITKRR 177  
DB 124 -----RRLCLVCGDVASGFHYGASCEACKAFKRTIOGNIIEYTCBANNECEINKRR 176  
QY 178 KSCQACRFMKCLKVMKEGVRIDRVRGGRQYKRRIDAENSPYLNQLVQPAKKPY--- 234  
DB 177 KACQACRFQKCLLMGMLKEGVRIDRVRGGRQYKRR--NPVNSYQTMQLLYQNTTSLCD 234  
QY 235 NKIVSHLLVAEPEKIYAM-PDPTVPDSI----- 262  
DB 235 VKLEVINSYEPDALSVQTPPPQVHTTSTINDAASSSSGSIKLESSVVTPTNGTCIFQNN 294  
QY 263 -----KALTTCDLADRELVTIGWAKHIFGSTLSLADQMSLIQSAWMEILLGVVRS 317  
DB 295 NNDPNEILLVSLDIYDKELVSVIGWAKQIPGFIPLINDQMKLIQVSWAELLTLQLTRRS 354  
QY 318 LSFEDLVYADDYIMDEDQSKLAGLLDNNAILQLVKKYKMKLEKEEFVTLKATLAN 377  
DB 355 LPFNGKLCFATDVWMDDEHLAKGCGYTFEYHCVQIAQRMERISPRREYLLKALLANC 414  
QY 378 DSMHIEDVEAVQKIQDVLAHALQDY-----EAGQHMEDPRRAGKMLMTLPILRQSTK 430  
DB 415 DIL-LDDQSSLRARFDTILNSLNDVVYLLRHSSAVSHQ-----QLLLLPRLQADDI 467  
QY 431 AVQHFYNIKLEKVPMEKLFLEMLE 455  
DB 468 LRRFWRGJARDEVITMKKLFLEMLE 492

RESULT 10  
Q8UWB0 PRELIMINARY; PRT; 587 AA.  
AC Q8UWB0;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Estrogen receptor.  
GN ER.

OS Caiman crocodilus (Spectacled caiman) (Caiman sclerops).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Crocodylidae; Alligatorinae; Caiman.  
OX NCBI\_TaxID=8499;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21490797; PubMed=11604222;  
RA Sumida K., Ooe N., Saito K., Kaneko H.;  
RT "Molecular cloning and characterization of reptilian estrogen receptor  
RT cDNAs";  
RL Mol. Cell. Endocrinol. 183:33-39(2001).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
DR EMBL; AB055220; BAB79436.1; -  
DR GO; GO:0016585; C:chromatin remodeling complex; ISS.  
DR GO; GO:0016020; C:membrane; ISS.  
DR GO; GO:0030284; F:estrogen receptor activity; ISS.  
DR GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.  
DR GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.  
DR InterPro; IPR000536; Hormone\_rec\_lig.  
DR InterPro; IPR001292; Oestrgn\_receptor.  
DR InterPro; IPR001723; Steroid\_receptor.  
DR InterPro; IPR008946; Steroid\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF02159; Oest\_recep; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDEFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;  
KW Transcription regulation; Zinc; Zinc-finger.  
SQ SEQUENCE 587 AA; 66514 MW; DE17F2B837FCF513 CRC64;

Query Match 30.4%; Score 726; DB 13; Length 587;  
Best Local Similarity 38.2%; Pred. No. 9.2e-52;  
Matches 175; Conservative 80; Mismatches 143; Indels 60; Gaps 14;

QY 38 FIKTEPSSPASLTDSVNHSPG-----SSDASGSYSSTMNQHONGDSPPLPSAPILGG 93  
DB 104 FLQTPA-----QLSPFVHHHSQVYLYLNDQSGF-----GMREAASTFYRPSADSRHQ 153  
QY 94 SGPVRKLYDDCSSTIVEDPQTKCEYMLNSMPK-RLCLVCGDIASGYHYGASCEACKAF 152  
DB 154 SGRER-----MSST-----SEKASLSMESTKETRYCAVCNDYASGYHYGWSCEGKAF 203  
QY 153 KRTIOGNIIEYSCPATNECEITKRRKSCQACRFMKCLKVMKEGVRIDRVRGGRQ-KYK 211  
DB 204 KRSIQGHNDYMCFAINQCTIDKNRRKSCQACRLRKCYEYGMKKGIRKDR-RGGRMLKOK 262  
QY 212 RR--IDAENSP-----YLNQLVQPAKK-----PYNKIVSHLLVAEPEKIY 250  
DB 263 KQREQDARNGETATTAEMRPTLTMTSPLVIKHTKKNSPALSLTAEQVSALEHAEPPIVY 322  
QY 251 AMPDPTVPDSIDIKALTTICDLADRELVTIGWAKHIFGSTLSLADQMSLIQSAWMEILI 310  
DB 323 SEYDPNRPFNFEASMTLLTLNLADRELVTIMWAKKRVPGFVDLTILHDQVHLDECAWLEILM 382  
QY 311 LGVYRSLSFEDLVYADDYIMDEDQSK-LAGLLDNNAILQLVKKYKMKLEKEEFVTL 369  
DB 383 IGLVWRSMHEPGKLLFAPNLLDRNQKCYEGWVEIFDMLLATARFRNMNLQGEHFVCL 442  
QY 370 KATALAN-----SDSMHIEDVEAVQKIQDVLAHALQDYEAGQHM--EDPRRAGKML 418  
DB 443 KSIILLNGSVYTFPLSTLKSLEEKDYIHRVLDKITDTLILHMAKSGSLQOQHRRLAQLL 502  
QY 419 MTLPLRLQSTKAVQHFYNIKLEKVPMEKLFLEMLEA 456  
DB 503 LILSHIRHMSNKGMEHLYNMCKKNVVPPLYDLLLEMLDA 540

RESULT 11  
Q8AYH0 PRELIMINARY; PRT; 589 AA.  
ID Q8AYH0  
AC Q8AYH0;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Estrogen receptor alpha.  
OS Coturnix coturnix japonica (Japanese quail).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Coturnix.  
OC NCBI\_TaxID=93934;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Liver;  
RC Ichikawa K., Yamamoto I., Tsukada A., Saito N., Shimada K.;  
RA "cDNA cloning and mRNA expression of estrogen receptor alpha in  
RT Japanese quail."  
RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF442965; AAN63674.1; -  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0005496; F:steroid binding; IEA.  
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR000536; Hormone\_rec\_lig.  
DR InterPro; IPR001292; Oestrgn\_receptor.  
DR InterPro; IPR001723; Stdhrmn\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF02159; Oest\_recep; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOLI; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR Receptor; PS00031; NUCLEAR\_RECEPTOR; 1.  
KW Receptor.  
SQ SEQUENCE 589 AA; 66767 MW; 8271FDA67552CA4 CRC64;  
Query Match 30.3%; Score 724; DB 13; Length 589;  
Best Local Similarity 36.6%; Pred. No. 1.4e-51;  
Matches 169; Conservative 89; Mismatches 136; Indels 68; Gaps 13;  
QY 38 FIKTEPSSPASLTDSVNHHS--PGSSSDASGSYSTMNGHONGLSPLY--PSAPILG 92  
DB 104 FLQTAPE---QLSPFIHHSQQVPYYLENEQGSFGMRET-----APPAFYRPS-- 148  
QY 93 GSGPVRKLYDDCSSTIVEDPQTKCE---YMLNSMPKRLCLVCGDIAAGYHYGVASCEAC 148  
DB 149 -----DNRHSIRERMSASSEKSGLSMESTKETRYCAVCNDYASGYHYGVWSCEGC 199  
QY 149 KAFPKRTIOGNIIEYSCPATNECEITKRRKSCQACRFMKCLKVGMLEKGVRLDRVGR- 207  
DB 200 KAFFKRSIQGNDYMCPATNQCTIDKNRKSQACRLRKCYEVGMKGIRKDR-RGGRM 258  
QY 208 QKYKRRIDAENSP-----YLNQLVQPAKK-----PYNKIVSHLLVAEP 246  
DB 259 MKQKRQREBQESRNGEASSTELRAPTLWTSPLYVKNKNSPALSLTAEGWVSALLLEAEP 318  
QY 247 EKIYAMPDPTVPDSIKALTTLCDLADRELVLIGWAKHIGFSTLSLADQMSLLQSAWM 306  
DB 319 PIYVSEYDPNRPFNEASMTLLTNLADRELVHMIMNAKRVPGFVDITLHDQVHLLCAWL 378  
QY 307 ELLILGVVYRSLFEDELIVYADDYIMEDQSK-LAGLLDLNNAIILQLVKKYKSMKLEKE 365  
DB 379 EILMIGLVWRSMHHPGKLLFAPNLLDRNQGKVEGVVHIFDMLLATARFRMMNLQGE 438

QY 366 FVTLKATAIAN-----SDSMHIEDYEAQOKLQDVLHEALQDYBAGQHM--EDPRRA 414  
DB 439 FVCLKSTILNSGVYTFLSSTLKSLEERYIHRVLDKITDTLHFMAKSGLSLQQQHRRL 498  
QY 415 GKMLMTPLRLQSTKAVQHFYNIKLEGVPMHKLFLMLEA 456  
DB 499 AOLLLILSHIRMSNKGMEHLYNMCKKNVPLYDLLLEMLDA 540  
RESULT 12  
Q8UWA9 PRELIMINARY; PRT; 581 AA.  
ID Q8UWA9  
AC Q8UWA9;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Estrogen receptor.  
GN ER.  
OS Chemidophorus uniparens (Desert grassland whiptail lizard).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Teiioidea;  
OC Teiidae; Chemidophorus.  
OC NCBI\_TaxID=37197;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=21490797; PubMed=11604222;  
RX Sumida K., Ooe N., Saito K., Kaneko H.;  
RA "Molecular cloning and characterization of reptilian estrogen receptor  
RT cDNAs."  
RT Mol. Cell. Endocrinol. 183:33-39(2001).  
RL -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
CC EMBL; AB055221; BAB79437.1; -  
DR GO; GO:0016585; C:chromatin remodeling complex; ISS.  
DR GO; GO:0016020; C:membrane; ISS.  
DR GO; GO:0030284; F:estrogen receptor activity; ISS.  
DR GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.  
DR GO; GO:0030520; F:estrogen receptor signaling pathway; ISS.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.  
DR InterPro; IPR000536; Hormone\_rec\_lig.  
DR InterPro; IPR001292; Oestrgn\_receptor.  
DR InterPro; IPR001723; Stdhrmn\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF02159; Oest\_recep; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOLI; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;  
KW Transcription regulation; Zinc; Zinc-finger.  
SQ SEQUENCE 581 AA; 65650 MW; A99F6C644FCF0D58 CRC64;  
Query Match 29.9%; Score 715; DB 13; Length 581;  
Best Local Similarity 34.9%; Pred. No. 7.5e-51;  
Matches 164; Conservative 80; Mismatches 133; Indels 93; Gaps 10;  
QY 30 HIDSSGSFIKTEPSSPASLT-----DSVNHSPGSSSDASGSYSTMNGHONGL 79  
DB 117 HHNQVPYYLENEPSSSAMREAFPTAFYRPGSENRHHGGRASNSEKGSLS----- 166  
QY 80 DSPPLYPAPILGSGPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIAAGYH 139  
DB 167 -----MESTKETRYCAVCNDYASGYH 187  
QY 140 YGVASCEACKAFPKRTIOGNIIEYSCPATNECEITKRRKSCQACRFMKCLKVGMLEKGV 199  
DB 188 YGVWSCEGCKAFKFKRSIQGNDYMCPATNQCTIDKNRKSQACRLRKCYEVGMKGIR 247

QY 200 LDRVRGRQ-KYKRRIID-----AENSP-YLNPQLVQPAKK-----PYNKIV 238  
Db 248 KDR-RGGMLKHKRQRBDELDERNAVAVTEARNTTLMPSPLMIKHSKNSPALSLTAEQMV 306  
QY 239 SHLVAEPEKIYAMPDPTVPDSIDIKALTTLCDLADRELVIIGWAKHIPGFSTLSLADQM 298  
Db 307 SALDAEPPIVSEYDPPSPFSEASVMTLLTNLADRELVMITWAKRVGFVDLALHDQV 366  
QY 299 SLIQSAMTEILIGVYRSLSFEDELVAADDYIMDEQSK-LAGLIDLNNAILQLVKKYK 357  
Db 367 HLEECAMTEILIMIGLWRSLEHFGKLLFAPNLLDRSQGMCVEGFVEIFDMLLATSSRFR 426  
QY 358 SMKLEKEEFVTLKAIALAN-----SDSMHIEDVAVQKLQDVLHEALQDYEAGQHM- 408  
Db 427 MNNIQGEFVCLKSIILLNSGYTFUSSTLRSLKEKEHHRVLDKITDTLTHMAKSGLS 486  
QY 409 --EDPRRAGKMLMTPLRLQSTFKAVQHFYNIKLEGVPMHKLFLMLEEA 456  
Db 487 LQQQHRRLAQLLMLSHIRHMSNKGMEHLYNMCKKNVPLDYDLLLEMLDA 536

RESULT 13

Q8QG80 PRELIMINARY; PRT; 152 AA.  
AC Q8QG80;  
DT 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-OCT-2002 (Tremblrel. 21, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Estrogen receptor-related receptor (Fragment).  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Westerlund L., Olsson P.-E.;  
RT "Cloning of zebrafish nuclear hormone receptors.";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
DR EMBL; AF495875; AAM18522.1; -.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR InterPro; IPR001723; Stdhmm\_receptor.  
DR InterPro; IPR008946; Str\_ncl\_receptor.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00398; STRDHOMONER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00399; Znf\_C4; 1.  
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;  
KW Transcription regulation; Zinc; Zinc-finger.  
FT NON\_TER 1  
FT NON\_TER 152  
SQ SEQUENCE 152 AA; 17189 MW; A97FA1BDABE091B1 CRC64;

Query Match 29.2%; Score 697.5; DB 13; Length 152;  
Best Local Similarity 87.4%; Pred. No. 3.1e-50;  
Matches 132; Conservative 5; Mismatches 7; Indels 7; Gaps 1;

QY 146 EACKAFFKRTIOGNIEXSCPATNCEITKRRRKSQAACRFMKCLKVGMLEKGVRLDRVVG 205  
Db 2 KGCKAFFKRTIOGNIEXSCPATNCEITKRRRKSQAACRFMKCLTVGMREGVRLDRVVG 61  
QY 206 GROKYYRRIDAENSPYLNQVLQPAKKPY-----NKIVSHLVAEPEKIYAMPDPTVP 258  
Db 62 GROKYYRRIDAENSPYLNQVLALPRKKPFPFGCLADNKIVSHLVAEPEKIYAMPDPTVP 121  
QY 259 DSDIKALTTLCDLADRELVIIGWAKHIPGF 289

Db 122 DSDIKALTTLCDLADRELVIIGWAKVPGF 152

RESULT 14

Q7T2K8 PRELIMINARY; PRT; 574 AA.  
AC Q7T2K8;  
DT 01-OCT-2003 (Tremblrel. 25, Created)  
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Estrogen receptor alpha.  
OS Halichoeres tenuispinnis.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;  
OC Labridae; Halichoeres.  
OX NCBI\_TaxID=217853;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Park J.-G., Kim S.-J.;  
RT "Molecular cloning and expression of two estrogen receptor subtypes in  
prologynous wrasse, Halichoeres tenuispinnis.";  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY305026; AAP72178.1; -.  
KW Receptor.  
SQ SEQUENCE 574 AA; 63661 MW; 34626475977971B1 CRC64;

Query Match 29.1%; Score 695; DB 13; Length 574;  
Best Local Similarity 33.9%; Pred. No. 3.4e-49;  
Matches 168; Conservative 90; Mismatches 169; Indels 68; Gaps 12;

QY 9 PESFSLHYEEELLCRMNKKDRHIDSSCSSFIKTEPSSPASLTDVNHHSFGSSDASGSX 68  
Db 31 PTLYSLSTQGYSAADTHGQPSDSIQS-LGSGPSSPLVFVPSSPRLSF-----F 80  
QY 69 SSTMGHQNGLDSPPLPSAPILGSGSPVRKLY--DDCSSTIVED---PQTKCEYMLNS 122  
Db 81 MHLPSHHYLETSTPYRRSSVSSSQSISRBEHCGTSDSYSMGEAGAAAGCEMAXE 140  
QY 123 MPKRLCLVCGDIASGYHYGVASCEACKAFAFKRTIOGNIEXSCPATNCEITKRRRKSQA 182  
Db 141 M--RYCAVCSDYASGYHYGVWSCEGCKAFAFKRSIQGHNDYMCPATNOCTIDNRNRKSCQA 198  
QY 183 CRFMKCLKVGMLEKGVRLDRV-----GROKYYKRI 214  
Db 199 CRLKCYEVGMMKGVKRDGRVLRDKRRRTGTSKXNGSKDRQRTVPQGRKRGSSV 258  
QY 215 DAENSPYLNQVLQPAKKPYNKIVSHLVAEPEKIYAMPDPTVPDSIDIKALTTLCDLADR 274  
Db 259 GGGKSPVIS-----MRPDQVLILLGAEPRLICSRQKLSRPYTEVTMTLLTSMTDR 310  
QY 275 ELVVIIGWAKHIPGFSTLSLADQMSLQSAMTEILIGVYRSLSFEDELVAADDYIMDE 334  
Db 311 ELVHMIAWAKKLPGFLQTLHDQVLESSWLEVMIGLWRSIHCPGKLIFAQDLIDR 370  
QY 335 DQSK-LAGLIDLNNAILQLVKKYKSMKLEKEFVTLKAIALANSDSMH-----IEDV 385  
Db 371 SEGDCEVGAETFDMLLATTSRFRMLKLPBEEFVCLKAIIILNSGAFSFGTGMPEPLHDN 430  
QY 386 EAVQKLQDVLHEALQDY--EAG--QHMEDDRAGKMLMTPLRLQSTFKAVQHFYNIKLE 441  
Db 431 EAVQNMLDITDALIHHSISQSGSAHQGS-RQAQLLLLSHTRHMSNKGMEHLYSMCK 489  
QY 442 GKVPMHKLFLMLEEA 456  
Db 490 NKVPLDYDLLLEMLDA 504

RESULT 15  
Q9UNJ4 PRELIMINARY; PRT; 129 AA.  
ID Q9UNJ4



AC Q9UNJ4;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Estrogen receptor related protein 3 (Fragment).  
GN ESR3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99357798; PubMed=10428842;  
RA Hong H., Yang L., Stallcup M.R.;  
RT "Hormone-independent transcriptional activation and coactivator  
binding by novel orphan nuclear receptor ESR3.";  
RL J. Biol. Chem. 274:22618-22626(1999).  
DR EMBL: AF117255; AAD48370.1; -;  
DR GO; GO:0004872; F:receptor activity; IEA.  
KW Receptor.  
FT NON TER  
SQ SEQUENCE 129 AA; 13990 MW; 7BE487F716CCD48 CRC64;  
Query Match 28.9%; Score 691; DB 4; length 129;  
Best Local Similarity 100.0%; Pred. No. 8.6e-50;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDSVELCLPESFSLHYEEELLCRMNSKDRHIDSSCSFIFKTEPSSPASLTDSVNHSPGG 60  
Db 1 MDSVELCLPESFSLHYEEELLCRMNSKDRHIDSSCSFIFKTEPSSPASLTDSVNHSPGG 60  
QY 61 SSDASGSYSTMNGHQNGLDSPLYPSAPILGSGPVRKLYDDCSSTIVEDPQTKCEYML 120  
Db 61 SSDASGSYSTMNGHQNGLDSPLYPSAPILGSGPVRKLYDDCSSTIVEDPQTKCEYML 120  
QY 121 NSMPKRLCL 129  
Db 121 NSMPKRLCL 129

Search completed: August 3, 2004, 12:41:49  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:39:31 ; Search time 19 Seconds

(without alignments)  
1244.458 Million cell updates/sec

Title: US-10-054-841-4

Perfect score: 2388

Sequence: 1 MDSVELCLPESFSLHYEEL.....KLEGVPMHKLFLMLEAKV 458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2388	100.0	458	3	US-09-141-000-4
2	2265	94.8	435	3	US-09-040-508-2
3	2265	94.8	435	4	US-09-500-654-2
4	2139	89.6	418	3	US-09-141-000-6
5	1799	75.3	500	3	US-09-141-000-2
6	1775	74.3	431	2	US-08-836-620A-19
7	1212.5	50.8	518	2	US-08-836-620A-18
8	716.5	30.0	596	2	US-08-836-620A-16
9	704	29.5	595	3	US-08-764-870-12
10	704	29.5	595	3	US-08-980-115-12
11	702	29.4	591	2	US-08-836-620A-17
12	702	29.4	595	3	US-09-041-886-35
13	702	29.4	595	4	US-08-453-998-2
14	685	28.7	410	6	5223606-5
15	661	27.7	484	2	US-08-836-620A-14
16	661	27.7	485	2	US-08-836-620A-5
17	655	27.4	548	3	US-09-139-617-1
18	655	27.4	548	4	US-09-561-741A-1
19	655	27.4	548	4	US-09-558-795-1
20	651	27.3	484	2	US-08-836-620A-13
21	651	27.3	485	2	US-08-836-620A-2
22	648	27.1	530	4	US-09-608-088-25
23	644.5	27.0	477	4	US-09-608-088-5
24	644.5	27.0	485	2	US-08-836-620A-3
25	597	25.0	111	3	US-09-040-508-4
26	597	25.0	111	4	US-09-500-654-4
27	596.5	25.0	467	1	US-08-336-408B-4

28	596.5	25.0	467	4	US-08-216-592A-6	Sequence 6, Appli
29	596.5	25.0	467	5	PCT-US91-00399-4	Sequence 4, Appli
30	589	24.7	462	1	US-08-336-408B-2	Sequence 2, Appli
31	589	24.7	462	3	US-08-764-870-6	Sequence 6, Appli
32	589	24.7	462	3	US-08-980-115-6	Sequence 2, Appli
33	589	24.7	462	5	PCT-US91-00399-2	Sequence 4, Appli
34	577.5	24.2	463	4	US-08-216-592A-8	Sequence 8, Appli
35	574	24.0	533	4	US-08-216-592A-4	Sequence 15, Appli
36	573.5	24.0	384	2	US-08-836-620A-15	Sequence 6, Appli
37	573.5	24.0	416	4	US-09-608-088-6	Sequence 21, Appli
38	573.5	24.0	418	4	US-07-952-800-2	Sequence 2, Appli
39	573	24.0	533	1	US-07-952-800-2	Sequence 6, Appli
40	572.5	24.0	463	1	US-08-336-408B-6	Sequence 6, Appli
41	572.5	24.0	463	5	PCT-US91-00399-6	Sequence 7, Appli
42	571	23.9	525	3	US-08-764-870-7	Sequence 4, Appli
43	571	23.9	525	3	US-08-980-115-7	Sequence 4, Appli
44	560	23.5	446	1	US-07-952-800-4	Sequence 2, Appli
45	558	23.4	448	4	US-08-216-592A-2	

ALIGNMENTS

RESULT 1  
US-09-141-000-4  
; Sequence 4, Application US/09141000  
; Patent No. 6054295  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Fang  
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
; FILE REFERENCE: 19999Y  
; CURRENT FILING DATE: 1998-08-26  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Human  
US-09-141-000-4

Query Match	100.0%;	Score 2388;	DB 3;	Length 458;
Best Local Similarity	100.0%;	Pred. No. 3.5e-238;		
Matches 458;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDSVELCLPESFSLHYEELLCRMSNKRHIDSSCSFFIKTEPSSPASLTDVNHSPGG	60	
DB	1	MDSVELCLPESFSLHYEELLCRMSNKRHIDSSCSFFIKTEPSSPASLTDVNHSPGG	60	
QY	61	SSDASGYSSTNGHQNGLDSPPLYPSPAPITIGSGPVRKLYDDCSSITVEDPQTKCEYML	120	
DB	61	SSDASGYSSTNGHQNGLDSPPLYPSPAPITIGSGPVRKLYDDCSSITVEDPQTKCEYML	120	
QY	121	NSMPKRLCLVCGDIAASGYHYGASCEACKAFKRTIOGNIIEYSCPATNECITKRRKSC	180	
DB	121	NSMPKRLCLVCGDIAASGYHYGASCEACKAFKRTIOGNIIEYSCPATNECITKRRKSC	180	
QY	181	QACRFMKCLKYGMLKEGVRLLDRVRGGRQYKRRIDAENSPYLPQLVOPAKKPYNKIVSH	240	
DB	181	QACRFMKCLKYGMLKEGVRLLDRVRGGRQYKRRIDAENSPYLPQLVOPAKKPYNKIVSH	240	
QY	241	LLVAEPEKIYAMDPTVPDSDIKALITLIDLADRELVTIIGWAKHIFGSTLSLADQMSL	300	
DB	241	LLVAEPEKIYAMDPTVPDSDIKALITLIDLADRELVTIIGWAKHIFGSTLSLADQMSL	300	
QY	301	LQSAWMEILLIGVYRSLSEFEDLVYADYIMDEDQSLAGLLDLNNAIQLVKKYKSMK	360	
DB	301	LQSAWMEILLIGVYRSLSEFEDLVYADYIMDEDQSLAGLLDLNNAIQLVKKYKSMK	360	
QY	361	LEKEEFVTLKAIALANSDSMHIIEDEVAVQKLQDVLHEALQDYEAGQHMEDPRRAGKMLMT	420	
DB	361	LEKEEFVTLKAIALANSDSMHIIEDEVAVQKLQDVLHEALQDYEAGQHMEDPRRAGKMLMT	420	

QY 421 LPLRQSTKAVQHFNKLEGVPMHKLFLMLEAKV 458  
DB 421 LPLRQSTKAVQHFNKLEGVPMHKLFLMLEAKV 458

RESULT 2

US-09-040-508-2  
; Sequence 2, Application US/09040508  
; Patent No. 6069239  
; GENERAL INFORMATION:  
; APPLICANT: Mathias, Stephen L.  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RATNER & PRESTIA  
; STREET: P.O. BOX 980  
; CITY: VALLEY FORGE  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19482  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/040,508  
; FILING DATE: 17-MAR-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9705451.4  
; FILING DATE: 17-MAR-1997  
; APPLICATION NUMBER: GB 9803289.9  
; FILING DATE: 16-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PRESTIA, PAUL F  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GH-30008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 435 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-040-508-2

Query Match 94.8%; Score 2265; DB 3; Length 435;  
Best Local Similarity 100.0%; Pred. No. 1.7e-225;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MSNKDRHIDSSCSSFIKTEPSSPASLTDSVNHSPGSSSDASGSYSTMNGHONGIDSP 83  
DB 1 MSNKDRHIDSSCSSFIKTEPSSPASLTDSVNHSPGSSSDASGSYSTMNGHONGIDSP 60  
QY 84 LYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKECEYMLNSMPKRLCLVCGDIASGYHYGVA 143  
DB 61 LYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKECEYMLNSMPKRLCLVCGDIASGYHYGVA 120  
QY 144 SCEACKAFKRTIOGNIETYSCPATNECEITKRRRKSQAACRFMKCLKVGMKKEGYRLDRV 203  
DB 121 SCEACKAFKRTIOGNIETYSCPATNECEITKRRRKSQAACRFMKCLKVGMKKEGYRLDRV 180  
QY 204 RGGROKYKRRIDAENSPYLNQVLQVPAKKPYNKIVSHLLVAEPEKIYAMPDPYVDSDIK 263  
DB 181 RGGROKYKRRIDAENSPYLNQVLQVPAKKPYNKIVSHLLVAEPEKIYAMPDPYVDSDIK 240  
QY 264 ALTTLCDLADRELVIIGWAKHIIPEFSTLSLADQMSLIQSAWMEIILIGVVYRSLSFED 323

DB 241 ALTTLCDLADRELVIIGWAKHIIPEFSTLSLADQMSLIQSAWMEIILIGVVYRSLSFED 300  
QY 324 LVIYADDYIMDEDQSLAGLIDLNNAILQLVKKYSMKLEKEEFVTLKALALANDSMHIE 383  
DB 301 LVIYADDYIMDEDQSLAGLIDLNNAILQLVKKYSMKLEKEEFVTLKALALANDSMHIE 360  
QY 384 DVEAVOKLQDVLHEALQDYEAGQHMEDPRRAGKMLMTLPLLRQSTKAVQHFNKLEGV 443  
DB 361 DVEAVOKLQDVLHEALQDYEAGQHMEDPRRAGKMLMTLPLLRQSTKAVQHFNKLEGV 420  
QY 444 VPMHKLFLMLEAKV 458  
DB 421 VPMHKLFLMLEAKV 435

RESULT 3

US-09-500-654-2  
; Sequence 2, Application US/09500654  
; Patent No. 6359116  
; GENERAL INFORMATION:  
; APPLICANT: Stephen L. Mathias  
; TITLE OF INVENTION: No. 6359116el Compounds  
; FILE REFERENCE: GP-30008-D1  
; CURRENT APPLICATION NUMBER: US/09/500,654  
; EARLIER FILING DATE: 2000-02-09  
; EARLIER FILING DATE: 1997-03-17  
; EARLIER FILING DATE: 1998-02-16  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 435  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
; US-09-500-654-2

Query Match 94.8%; Score 2265; DB 4; Length 435;  
Best Local Similarity 100.0%; Pred. No. 1.7e-225;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MSNKDRHIDSSCSSFIKTEPSSPASLTDSVNHSPGSSSDASGSYSTMNGHONGIDSP 83  
DB 1 MSNKDRHIDSSCSSFIKTEPSSPASLTDSVNHSPGSSSDASGSYSTMNGHONGIDSP 60  
QY 84 LYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKECEYMLNSMPKRLCLVCGDIASGYHYGVA 143  
DB 61 LYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKECEYMLNSMPKRLCLVCGDIASGYHYGVA 120  
QY 144 SCEACKAFKRTIOGNIETYSCPATNECEITKRRRKSQAACRFMKCLKVGMKKEGYRLDRV 203  
DB 121 SCEACKAFKRTIOGNIETYSCPATNECEITKRRRKSQAACRFMKCLKVGMKKEGYRLDRV 180  
QY 204 RGGROKYKRRIDAENSPYLNQVLQVPAKKPYNKIVSHLLVAEPEKIYAMPDPYVDSDIK 263  
DB 181 RGGROKYKRRIDAENSPYLNQVLQVPAKKPYNKIVSHLLVAEPEKIYAMPDPYVDSDIK 240  
QY 264 ALTTLCDLADRELVIIGWAKHIIPEFSTLSLADQMSLIQSAWMEIILIGVVYRSLSFED 323  
DB 241 ALTTLCDLADRELVIIGWAKHIIPEFSTLSLADQMSLIQSAWMEIILIGVVYRSLSFED 300  
QY 324 LVIYADDYIMDEDQSLAGLIDLNNAILQLVKKYSMKLEKEEFVTLKALALANDSMHIE 383  
DB 301 LVIYADDYIMDEDQSLAGLIDLNNAILQLVKKYSMKLEKEEFVTLKALALANDSMHIE 360  
QY 384 DVEAVOKLQDVLHEALQDYEAGQHMEDPRRAGKMLMTLPLLRQSTKAVQHFNKLEGV 443  
DB 361 DVEAVOKLQDVLHEALQDYEAGQHMEDPRRAGKMLMTLPLLRQSTKAVQHFNKLEGV 420  
QY 444 VPMHKLFLMLEAKV 458  
DB 421 VPMHKLFLMLEAKV 435



RESULT 4  
US-09-141-000-6  
; Sequence 6, Application US/09141000  
; Patent No. 6054295  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Fang  
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
; FILE REFERENCE: 19999Y  
; CURRENT APPLICATION NUMBER: US/09/141,000  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Human  
US-09-141-000-6

Query Match 89.6%; Score 2139; DB 3; Length 418;  
Best Local Similarity 100.0%; Pred. No. 1.8e-212;  
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDSVELCLPEPSFLHYEELLCRMSNKRHDSSCSFIKTEPSSPASITDSVNHSPGG 60  
Db 1 MDSVELCLPEPSFLHYEELLCRMSNKRHDSSCSFIKTEPSSPASITDSVNHSPGG 60  
QY 61 SSDASGSYSTNMGHONGLDSPPLYPSPALIGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120  
Db 61 SSDASGSYSTNMGHONGLDSPPLYPSPALIGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120  
QY 121 NSMPKRLCLVCGDIASGYHYGVASCACAFPKRTIQGNIIEYSCPATNECEITKRRKSC 180  
Db 121 NSMPKRLCLVCGDIASGYHYGVASCACAFPKRTIQGNIIEYSCPATNECEITKRRKSC 180  
QY 181 QACRFMKCLKVGMLEKGVRLDRVRGGROKYKRRIDAENSPLYNPQLVQPAKKPYNKIVSH 240  
Db 181 QACRFMKCLKVGMLEKGVRLDRVRGGROKYKRRIDAENSPLYNPQLVQPAKKPYNKIVSH 240  
QY 241 LIVAEPKITYAMPDPTVPDSIDIKALTTCDLADRELVIIGWAKHIFGSTLSLADQMSL 300  
Db 241 LIVAEPKITYAMPDPTVPDSIDIKALTTCDLADRELVIIGWAKHIFGSTLSLADQMSL 300  
QY 301 LOSAMWEILLGVYRSLSEFDELIVADYIMDEDQSKLAGLLDNNAILQLVKYKYSMK 360  
Db 301 LOSAMWEILLGVYRSLSEFDELIVADYIMDEDQSKLAGLLDNNAILQLVKYKYSMK 360  
QY 361 LEKEEFVTLKAILANSDSMHIEDVEAVQKLQDVLHEALQDYEAQOME 409  
Db 361 LEKEEFVTLKAILANSDSMHIEDVEAVQKLQDVLHEALQDYEAQOME 409

RESULT 5  
US-09-141-000-2  
; Sequence 2, Application US/09141000  
; Patent No. 6054295  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Fang  
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
; FILE REFERENCE: 19999Y  
; CURRENT APPLICATION NUMBER: US/09/141,000  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Human  
US-09-141-000-2

Query Match 75.3%; Score 1799; DB 3; Length 500;  
Best Local Similarity 78.1%; Pred. No. 3.6e-177;  
Matches 339; Conservative 50; Mismatches 43; Indels 2; Gaps 2;  
QY 24 MSNKRHDSSCSFIKTEPSSPASITDSVNHSPGGSSDASGSYSTNMGHONGLDSP 83  
Db 1 MSSDDRHLGSSCGSFIKTEPSSPSGIDALSHSPSSGSSDASGFGALGTTHANGLDSP 60  
QY 84 LYPAPILGSGPYRKLVDCCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 143  
Db 61 MPAGAG-LGRT-PCKSYEDCAGIMEDSAIKCEYMLNAILPKRLCLVCGDIASGYHYGA 118  
QY 144 SCEACKAFPKRTIQGNIIEYSCPATNECEITKRRKSCQACRFMKCLKVGMLEKGVRLDRV 203  
Db 119 SCEACKAFPKRTIQGNIIEYSCPATNECEITKRRKSCQACRFMKCLKVGMLEKGVRLDRV 178  
QY 204 RGRQKXKRRIDAENSPLYNPQLVQPAKKPYNKIVSHLIVAEPKITYAMPDPTVPDSIDIK 263  
Db 179 RGRQKXKRRIDSESSPYLSLQISPPAKKPLTKIVSYLLVAEPDKLYAMPDPGMEGDIK 238  
QY 264 ALTTCLADRELVIIGWAKHIFGSTLSLADQMSLLOSAMWEILLGVYRSLSEFE 323  
Db 239 ALTTCLADRELVIIGWAKHIFGSTLSLADQMSLLOSAMWEILLGVYRSLSEFE 298  
QY 324 LVAADDYIMDEDQSKLAGLLDNNAILQLVKYKYSMKLEKEEFVTLKAILANSDSMHIE 383  
Db 299 LVAADDYIMDEDQSKLAGLLDNNAILQLVKYKYSMKLEKEEFVTLKAILANSDSMHIE 358  
QY 384 DVEAVQKLQDVLHEALQDYEAQOMEDEPRRAGKMLMTLPLRQTSKAVQHFYNIKLEK 443  
Db 359 DVEAVQKLQDVLHEALQDYELSRHEEPWRTKLLTLLPLRQTAARAVQHFYSVKLQK 418  
QY 444 VPMHKLFLMLEAK 457  
Db 419 VPMHKLFLMLEAK 432

RESULT 6  
US-08-836-620A-19  
; Sequence 19, Application US/08836620A  
; Patent No. 5958710  
; GENERAL INFORMATION:  
; APPLICANT: Orphan receptor  
; TITLE OF INVENTION: Orphan receptor  
; NUMBER OF SEQUENCES: 19  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,620A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP96/03933  
; FILING DATE:  
; APPLICATION NUMBER: GB 9518272.1  
; FILING DATE: 08-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9605550.4  
; FILING DATE: 15-MAR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9607532.0  
; FILING DATE: 11-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9609576.5  
; FILING DATE: 08-MAY-1996  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 431 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; ORIGINAL SOURCE:



; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 596 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; ORIGINAL SOURCE:  
; ORGANISM: Rattus rattus  
; US-08-836-620A-16

Query Match 30.0%; Score 716.5; DB 2; Length 596;  
Best Local Similarity 39.2%; Pred. No. 5.1e-65;  
Matches 168; Conservative 82; Mismatches 128; Indels 51; Gaps 13;

QY 73 NGHONGLDSPPLY-----PSAPILGGGP---VRKLYDDCSSTIVEDPQTKE----YMLN 121  
Db 129 HGHQ-----VPYLENEPSAYAVDTGPPAFYRSNSDNRRQNGRERLSSSEKGNMIMES 183  
QY 122 SMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIIEYSCPATNECEITKRRKGCQ 181  
Db 184 AKETRYCAVCNDYASGYHYGVSCGCKAFKRSIQGHNDYMCPATNQCTIDKNRRKSCQ 243  
QY 182 ACRFMKCLKVGMKEGVRLLDRVGGGRQ-KYKR-RIDAEN-----SPYLNPOL 226  
Db 244 ACRLRKCYEVGMKGIKDR-RGGRMLKHKRQDDLEGNNEMGTSGDMRAANLWSPPLY 302  
QY 227 VQPAKK-----PYNKIVSHLLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADRELVI 279  
Db 303 IKHTKKNSPALSTADQWVSALLDAEPPLIYSEYDPSRPFSEASMGILLTNLADRELVM 362  
QY 280 IGWAKHIPGFSTLSLADQMSLLQSAWMEIILGVVYRSLSFEDLVYADDYIMDEDSK- 338  
Db 363 INWAKRVPGFGLNLHQVHLLECAWLEIMIGLVWRSMHEHPGKLLFAPNLLDRNQKC 422  
QY 339 IAGLLDLNNAIQLVKKYKSMKLEKEFVTLKAIALAN-----SDSMTEDEAVQK 390  
Db 423 VEGWVEIFDMLLATSSRFMMNLQGEFFVCLKSIILNSGVYTFLLSTLKSLEKDHHR 482  
QY 391 LQDVLEALQDY--EAGQMEDP-RRAGKMLMTPLRLQSTSTKAVQHFYNIKLEKVPMH 447  
Db 483 VLDKINDTLIHMAKAGTLTQQOHRRLAQLLILSHIRMSNKGMEHLVXMKCNVVPPLY 542  
QY 448 KLFLEMLEA 456  
Db 543 DLLLEMLDA 551

RESULT 9

US-08-764-870-12  
; Sequence 12, Application US/08764870  
; Patent No. 6236946  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Thomas S  
; APPLICANT: Baxter, John D  
; APPLICANT: Fletterick, Robert J  
; APPLICANT: Wagner, Richard L  
; APPLICANT: Kushner, Peter J  
; APPLICANT: Apriletti, James W  
; TITLE OF INVENTION: Nuclear Receptor ligands and ligand  
; TITLE OF INVENTION: Binding Domains  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooley Godward  
; STREET: Five Palo Alto Square, 3000 El Camino Real  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/764,870  
; FILING DATE: 13-DEC-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/008,540  
; FILING DATE: 13-DEC-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/008,543  
; FILING DATE: 13-DEC-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/008,606  
; FILING DATE: 14-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Nakamura, Jackie N  
; REGISTRATION NUMBER: 35,966  
; REFERENCE/DOCKET NUMBER: UCAL-246/01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650)843-5000  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 595 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-764-870-12

Query Match 29.5%; Score 704; DB 3; Length 595;  
Best Local Similarity 41.2%; Pred. No. 1e-63;  
Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

QY 119 MLNSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIIEYSCPATNECEITKRRK 178  
Db 176 MESAKETRYCAVCNDYASGYHYGVSCGCKAFKRSIQGHNDYMCPATNQCTIDKNRRK 235  
QY 179 SCQACRFMKCLKVGMKEGVRLLDRVGGGRQ-KYKR-RIDAEN-----SPYLN 223  
Db 236 SCQACRLRKCYEVGMKGIKDR-RGGRMLKHKRQDDGEGRGEVGSAGDMRAANLWPS 294  
QY 224 POLVQPAKK-----PYNKIVSHLLVAEPEKIYAMPDPTVPDSDIKALTTLCDLADREL 276  
Db 295 PLMIKRSKXNSLALSTADQWVSALLDAEPPLIYSEYDPTRFSEASMGILLTNLADREL 354  
QY 277 VVIIGWAKHIPGFSTLSLADQMSLLQSAWMEIILGVVYRSLSFEDLVYADDYIMDEQ 336  
Db 355 VHMIMWAKRVPGFVDLTLDQVHLLECAWLEIMIGLVWRSMHEHPGKLLFAPNLLDRNQ 414  
QY 337 SK-IAGLLDLNNAIQLVKKYKSMKLEKEFVTLKAIALANS-----DSMH 381  
Db 415 GKQVEGWVEIFDMLLATSSRFMMNLQGEFFVCLKSIILNSGVYTFLLSTLKSLEKDH 474  
QY 382 IEDVEAVQKLQDVLEALQDYEAGQMEDP-RRAGKMLMTPLRLQSTSTKAVQHFYNIKL 440  
Db 475 IHRV--LDKITDTLIHMA--KAGTLTQQOHRRLAQLLILSHIRMSNKGMEHLYSMKC 530  
QY 441 EGVVPMHKLPLEMLEA 456  
Db 531 KVVVPLYDLLLEMLDA 546

RESULT 10

US-08-980-115-12  
; Sequence 12, Application US/08980115  
; Patent No. 6266622  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Thomas S.  
; APPLICANT: Baxter, John D.  
; APPLICANT: Fletterick, Robert J.  
; APPLICANT: Wagner, Richard L.  
; APPLICANT: Kushner, Peter J.  
; APPLICANT: Apriletti, James W.  
; APPLICANT: West, Brian L.



APPLICANT: Shiau, Andrew K.  
TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS  
FILE REFERENCE: UCAL-246/020US  
CURRENT APPLICATION NUMBER: US/08/980,115  
CURRENT FILING DATE: 1997-11-26  
EARLIER APPLICATION NUMBER: 08/764,870  
EARLIER FILING DATE: 1996-12-13  
EARLIER APPLICATION NUMBER: 60/008,606  
EARLIER FILING DATE: 1995-12-14  
EARLIER APPLICATION NUMBER: 60/008,543  
EARLIER FILING DATE: 1995-12-13  
EARLIER APPLICATION NUMBER: 60/008,540  
EARLIER FILING DATE: 1995-12-13  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12  
LENGTH: 595  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (287)..(549)  
OTHER INFORMATION: minimal ligand binding domain  
US-08-980-115-12

Query Match 29.5%; Score 704; DB 3; Length 595;  
Best Local Similarity 41.2%; Pred. No. 1e-63;  
Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

QY 119 MLNSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIYSCPATNECEITKRRK 178  
DB 176 MESAKETRYCAVCNDYASGYHYGVWSCGCKAFKRSIQGNDYMCPTNQCITDKNRK 235  
QY 179 SCQACRFMKCLKVGMLKEGVRLDRVGRGQ-KYKR-RIDAEN-----SPYLN 223  
DB 236 SCQACRLRKCYEVGMKGGIRKDR-RGGRMLKHKRQDDGEGRGVGSAGDMRANLWPS 294  
QY 224 POLVQPAK-----PYNKIVSHLLVAEPEKIYAMPDPTVPDSIDIKALTTLCIDLREL 276  
DB 295 PLMIKRSKNSLALSLTADQWVSALLDAEPILYSEYDPTRPFESEASMGILTINLADREL 354  
QY 277 VVIIGWAKHIPGFSTLSLADQMSLLQSAMWEILLGVYRSLSEFDELVYADDTYIMDEDQ 336  
DB 355 VHMIMWAKRVPGEFVDLTLDQVHLLBCAMLEIMIGLVWRSMHEHPKLLFAPNLLDRNQ 414  
QY 337 SK-LAGLLDNLNAILQLVKKYKSMKLEKEEFVTIKALALANS-----DSMH 381  
DB 415 GKCVGEWVEIFDMLLATSSRRPMNLQGEFEVCLKSILLNSGVYTFSLSTLKSLEKDH 474  
QY 382 IEDVEAVQKLQDVLHEALQDYEAQOMEDP-RRAGKMLMTPLLRQSTKAVQHFYNIKL 440  
DB 475 IHRV--LDKITDTLILHMA--KAGLTQQQHQRLAQLLLISHIRMSNKGMEHLYSMKC 530  
QY 441 EGKVPMEKLFLEMLEA 456  
DB 531 KNVVPPLYDLLEMLDA 546

RESULT 11  
US-08-836-620A-17  
Sequence 17, Application US/08836620A  
Patent No. 5958710  
GENERAL INFORMATION:  
APPLICANT: Orphan receptor  
TITLE OF INVENTION: Orphan receptor  
NUMBER OF SEQUENCES: 19  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/836,620A

FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP96/03933  
FILING DATE:  
APPLICATION NUMBER: GB 9518272.1  
FILING DATE: 08-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9605550.4  
FILING DATE: 15-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9607532.0  
FILING DATE: 11-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9609576.5  
FILING DATE: 08-MAY-1996  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 591 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-836-620A-17

Query Match 29.4%; Score 702; DB 2; Length 591;  
Best Local Similarity 41.2%; Pred. No. 1.6e-63;  
Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

QY 119 MLNSMPKRLCLVCGDIASGYHYGVASCEACKAFKRTIQGNIYSCPATNECEITKRRK 178  
DB 176 MESAKETRYCAVCNDYASGYHYGVWSCGCKAFKRSIQGNDYMCPTNQCITDKNRK 235  
QY 179 SCQACRFMKCLKVGMLKEGVRLDRVGRGQ-KYKR-RIDAEN-----SPYLN 223  
DB 236 SCQACRLRKCYEVGMKGGIRKDR-RGGRMLKHKRQDDGEGRGVGSAGDMRANLWPS 294  
QY 224 POLVQPAK-----PYNKIVSHLLVAEPEKIYAMPDPTVPDSIDIKALTTLCIDLREL 276  
DB 295 PLMIKRSKNSLALSLTADQWVSALLDAEPILYSEYDPTRPFESEASMGILTINLADREL 354  
QY 277 VVIIGWAKHIPGFSTLSLADQMSLLQSAMWEILLGVYRSLSEFDELVYADDTYIMDEDQ 336  
DB 355 VHMIMWAKRVPGEFVDLTLDQVHLLBCAMLEIMIGLVWRSMHEHPKLLFAPNLLDRNQ 414  
QY 337 SK-LAGLLDNLNAILQLVKKYKSMKLEKEEFVTIKALALANS-----DSMH 381  
DB 415 GKCVGEWVEIFDMLLATSSRRPMNLQGEFEVCLKSILLNSGVYTFSLSTLKSLEKDH 474  
QY 382 IEDVEAVQKLQDVLHEALQDYEAQOMEDP-RRAGKMLMTPLLRQSTKAVQHFYNIKL 440  
DB 475 IHRV--LDKITDTLILHMA--KAGLTQQQHQRLAQLLLISHIRMSNKGMEHLYSMKC 530  
QY 441 EGKVPMEKLFLEMLEA 456  
DB 531 KNVVPPLYDLLEMLDA 546

RESULT 12  
US-09-041-886-35  
Sequence 35, Application US/09041886  
Patent No. 6235872  
GENERAL INFORMATION:  
APPLICANT: Bredesen, Dale E.  
TITLE OF INVENTION: Proapoptotic Peptides, Dependence  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States

ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/041,886  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 2626  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 595 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-041-886-35

Query Match 29.4%; Score 702; DB 3; Length 595;  
Best Local Similarity 41.2%; Pred. No. 1.6e-63;  
Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

QY 119 MLNSMPKRLCLVCGDIASGYHYGVASCEACAFKRTIQGNIYSCPATNECEITKRRK 178  
DB 176 MESAKETRYCAVNDYASGYHYGVASCEGCKAFKRSIQGHNDYMCPTATNQTIDKNRRK 235  
QY 179 SCQACRFMKCLKVGMLKEGVRDLRVGGRQ-KYKR-RIDAEN-----SPYLN 223  
DB 236 SCQACRLRKCYEVGMKKGIRKDR-RGGRMLKHKRQRDDGEGRGVGSAGDMRAANLWPS 294  
QY 224 POLVOPAKK-----PYNKIYSHLLVAEPEKIYAMPDPTVPDSIDIKALITLCLADREL 276  
DB 295 PLMKRSKKNLSLSTLADQWVSALLDAEPPIYSEYDPTRFSEASMMGLTLNLADREL 354  
QY 277 VVIIGWAKHIPGFSTLSLADQMSLLQSAMEILLGVVYRSLSFEDLVYADDYIMDEDQ 336  
DB 355 VHMIMWAKRVPGFVDLTLDQVHLLBECAMLEILMIGLVWMSMEHPVKLLFAPNLLDRNQ 414  
QY 337 SK-LAGLLDLNNAILLQVKKYKSMKLEKEEFVTLKATIALANS-----DSMH 381  
DB 415 GKCVGEWVEIFDMLATSSRFPMNMLQGEFVCLKSIITLINSGVYTFLSSTLKSLERKH 474  
QY 382 IEDVEAVQKLQDVHLEALQDYEAQHMEDP-RRAGKMLMTLPILRQSTKAVQHFYNTKL 440  
DB 475 IHRV--LDKITDTLILHMA--KAGLTLLQOQHRLAQLLLISHIRHMSNKGMEHLYSMKC 530  
QY 441 EGKVPMKLFLMLEA 456  
DB 531 KNVVPPLYDLLLEMLDA 546

RESULT 13  
US-08-453-998-2  
Sequence 2, Application US/08453998  
Patent No. 6444438  
GENERAL INFORMATION:  
APPLICANT: CHAMON, PIERRE  
APPLICANT: METZGER, DANIEL  
APPLICANT: WHITE, JOHN  
TITLE OF INVENTION: METHOD FOR THE PREPARATION OF A PROTEIN  
TITLE OF INVENTION: BY YEASTS USING AN INDUCIBLE SYSTEM, VECTORS AND  
TITLE OF INVENTION: CORRESPONDING TRANSFORMED STRAINS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1100 NEW YORK AVE., N.W.

CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,998  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/161,064  
FILING DATE: 03-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: CHAPIN, MARIANA K.  
REGISTRATION NUMBER: 35,843  
REFERENCE/DOCKET NUMBER: 1037/98493  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3711  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 595 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-453-998-2

Query Match 29.4%; Score 702; DB 4; Length 595;  
Best Local Similarity 41.2%; Pred. No. 1.6e-63;  
Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

QY 119 MLNSMPKRLCLVCGDIASGYHYGVASCEACAFKRTIQGNIYSCPATNECEITKRRK 178  
DB 176 MESAKETRYCAVNDYASGYHYGVASCEGCKAFKRSIQGHNDYMCPTATNQTIDKNRRK 235  
QY 179 SCQACRFMKCLKVGMLKEGVRDLRVGGRQ-KYKR-RIDAEN-----SPYLN 223  
DB 236 SCQACRLRKCYEVGMKKGIRKDR-RGGRMLKHKRQRDDGEGRGVGSAGDMRAANLWPS 294  
QY 224 POLVOPAKK-----PYNKIYSHLLVAEPEKIYAMPDPTVPDSIDIKALITLCLADREL 276  
DB 295 PLMKRSKKNLSLSTLADQWVSALLDAEPPIYSEYDPTRFSEASMMGLTLNLADREL 354  
QY 277 VVIIGWAKHIPGFSTLSLADQMSLLQSAMEILLGVVYRSLSFEDLVYADDYIMDEDQ 336  
DB 355 VHMIMWAKRVPGFVDLTLDQVHLLBECAMLEILMIGLVWMSMEHPVKLLFAPNLLDRNQ 414  
QY 337 SK-LAGLLDLNNAILLQVKKYKSMKLEKEEFVTLKATIALANS-----DSMH 381  
DB 415 GKCVGEWVEIFDMLATSSRFPMNMLQGEFVCLKSIITLINSGVYTFLSSTLKSLERKH 474  
QY 382 IEDVEAVQKLQDVHLEALQDYEAQHMEDP-RRAGKMLMTLPILRQSTKAVQHFYNTKL 440  
DB 475 IHRV--LDKITDTLILHMA--KAGLTLLQOQHRLAQLLLISHIRHMSNKGMEHLYSMKC 530  
QY 441 EGKVPMKLFLMLEA 456  
DB 531 KNVVPPLYDLLLEMLDA 546

RESULT 14  
5223606-5  
Patent No. 5223606  
APPLICANT: BLAUDIN DE THE, HUGHES, MARCHIO, AGNES, TIOLLAIS,  
PIERRE, DEJEAN, ANNE  
TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED  
PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA

NUMBER OF SEQUENCES: 11  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/134,130  
FILING DATE: 17-DEC-1987  
PRIOR APPLICATION DATA:

SEQ ID NO:5:  
LENGTH: 410

5223606-5

Query Match 28.7%; Score 685; DB 6; length 410;  
Best Local Similarity 41.7%; Pred. No. 5e-62;  
Matches 153; Conservative 73; Mismatches 97; Indels 44; Gaps 11;

QY 128 CLVCGDIASGYHYGVASCEACAFKKRTIQGNIYSCPATNECEITKRRKSCQACRFMK 187  
DB 1 CAVCNDYASGYHYGVWSCEGCKAFKKRSIQGNDYMCPATNQCTIDKNRRKSCQACRLRK 60  
QY 188 CLKVGMLEKGVRLDRVRGGRQ-KYKR-RIDAEN-----SPYLNPLVQPAKK 232  
DB 61 CYEVGMKKGIRKDR-RGGRMLKHKRQDDGEGRGVGSAGDMRANLWPSPLMIKRSKK 119  
QY 233 -----PYNKIVSHLLVAEPEKIYAMPDETPVPSDIKALTTLCDLADRELVTIGMAKH 285  
DB 120 NSLALSITADQVWSALLDAEP-ILYSEYDPTRFSEASMMGLTNLADRELVHMINWAKR 178  
QY 286 IPGFSTLSIADQMSLQSAWMEILILGVVYRSLSFEDELVYADYIMDEDQSK-LAGLLD 344  
DB 179 VPGFVDLTLDQVHLIECAWLEILMIGLVRSMEHPVKLLFAPNLLDRNQSKVEGMVE 238  
QY 345 LNNAILQLVKYKSMKLEKEEFTVTKAIALANS-----DSMHIEDVEAVQK 390  
DB 239 IFDMLLATSRFRMNNLQGESEFVCLKSIILNSGVYTFLSSTLKSLEEDHIIHRV-LDK 296  
QY 391 LQDVLHEALQDVEAGQMEDP-RRAGKMLMTPLLRQSTKAVQHFYNIKLEGVPMHKL 449  
DB 297 ITDTLHLMA--KAGLTLQOQHQLAQLLLSHIRMSNKGMEHLYSMCKXNVVPLVDL 354  
QY 450 FLEMLEA 456  
DB 355 LLEMLDA 361

RESULT 15

US-08-836-620A-14  
Sequence 14, Application US/08836620A  
Patent No. 5958710

GENERAL INFORMATION:

APPLICANT:  
TITLE OF INVENTION: Orphan receptor  
NUMBER OF SEQUENCES: 19  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/836,620A  
FILING DATE:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP96/03933  
FILING DATE:

APPLICATION NUMBER: GB 9518272.1  
FILING DATE: 08-SEP-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9605550.4  
FILING DATE: 15-MAR-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9607532.0  
FILING DATE: 11-APR-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9609576.5  
FILING DATE: 08-MAY-1996

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
US-08-836-620A-14

Query Match 27.7%; Score 661; DB 2; length 484;  
Best Local Similarity 38.6%; Pred. No. 2e-59;  
Matches 150; Conservative 79; Mismatches 118; Indels 42; Gaps 11;

QY 91 LGSGPVRKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCIVCGDIASGYHYGVASCEACKA 150  
DB 83 LGSG-----CASP-VTSPSTKRD-----AHFCAYCSDYASGYHYGVWSCEGCKA 126  
QY 151 FFKRTIQGNIYSCPATNECEITKRRKSCQACRFMKCLKVGMLEKGVRLDRVRGGRQKY 210  
DB 127 FFKRSIQGNDYICPATNQCTIDKNRRKNCQACRLRKCYEVGMVWVCGSRRERC-GYRIYR 185  
QY 211 KRRIDAENSPYLN-----POLVQPAKKPYNKIVSHLLVAEPEKIYAMPDETPV 258  
DB 186 RQRSASEQVHCLNKAARTSGHTPRVKEILLNSLSP-EQLVLTLLAEAPPNV-LVSRPSMP 243  
QY 259 DSDIKALTTLCDLADRELVTIGMAKHIPGFSTLSIADQMSLQSAWMEILILGVVYRSI 318  
DB 244 FTEASMMWSLTKLADKELVHMIGWAKKIPGFVELSLDDQVRLLESQWMEVLMVGLMWSRI 303  
QY 319 SFEDELVYADYIMDEDQSK-LAGLLDNNAILQLVKYKSMKLEKEEFTVTKAIALANS 377  
DB 304 DHPGKLIFAPDLVLDREGEKVEGILEIFDMLLATTAFFRELKIQHKEYLCVKAMILLNS 363  
QY 378 DSMHI-----EDVEAVQKLQDVLH--EALQDVEAGQMEDPR--RAGKMLMTPLLRQT 427  
DB 364 SMYHLATASQEAHSSRKLTHLINAVTDALVWVISKRSISSQOQSVRLANLMLLSHVRI 423  
QY 428 STKAVQHFYNIKLEGVPMHKL FLEMLEA 456  
DB 424 SNKGMEHLLSMCKXNVVPPYDILLLEMLNA 452

Search completed: August 3, 2004, 12:42:46  
Job time : 20 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:40:36 ; Search time 45 Seconds  
(without alignments)  
3192.595 Million cell updates/sec

Title: US-10-054-841-4  
Perfect score: 2388  
Sequence: 1 MDSVELCLPESFSLHYEHEL.....KLEGGVPMHKLFLMLEAKV 458

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2388	100.0	458	14 US-10-054-841-4	Sequence 4, Appli
2	2388	100.0	458	15 US-10-355-218-1	Sequence 1, Appli
3	2139	89.6	418	14 US-10-054-841-6	Sequence 6, Appli
4	1799	75.3	500	14 US-10-054-841-2	Sequence 2, Appli
5	1775	74.3	431	14 US-10-278-481-19	Sequence 19, Appli
6	1259	52.7	486	15 US-10-297-653-1	Sequence 1, Appli
7	1258.5	52.7	519	14 US-10-211-239-1	Sequence 1, Appli
8	1212.5	50.8	518	14 US-10-278-481-18	Sequence 18, Appli
9	1073	44.9	516	15 US-10-093-463-202	Sequence 202, App
10	947	39.7	230	15 US-10-355-218-8	Sequence 8, Appli
11	755	31.6	234	15 US-10-355-218-9	Sequence 9, Appli
12	726	30.4	587	16 US-10-450-484A-1	Sequence 1, Appli
13	716.5	30.0	596	14 US-10-278-481-16	Sequence 16, Appli
14	715	29.9	581	16 US-10-451-768A-1	Sequence 1, Appli
15	713.5	29.9	595	14 US-10-052-092-31	Sequence 31, Appli

16	713.5	29.9	595	14 US-10-437-107-31	Sequence 31, Appli
17	711.5	29.8	595	14 US-10-052-092-30	Sequence 30, Appli
18	711.5	29.8	595	14 US-10-437-107-30	Sequence 30, Appli
19	707	29.6	595	15 US-10-148-835-8	Sequence 8, Appli
20	704	29.5	595	9 US-09-853-033-2	Sequence 2, Appli
21	704	29.5	595	15 US-10-148-835-1	Sequence 1, Appli
22	704	29.5	595	15 US-10-148-835-3	Sequence 3, Appli
23	704	29.5	595	15 US-10-148-835-9	Sequence 9, Appli
24	703	29.4	595	15 US-10-148-835-5	Sequence 5, Appli
25	702	29.4	591	14 US-10-278-481-17	Sequence 17, Appli
26	702	29.4	595	9 US-09-933-267A-2	Sequence 2, Appli
27	702	29.4	595	10 US-09-952-680A-10	Sequence 10, Appli
28	702	29.4	595	13 US-10-096-710-1	Sequence 1, Appli
29	702	29.4	595	13 US-10-081-563-2	Sequence 2, Appli
30	702	29.4	595	14 US-10-052-092-9	Sequence 9, Appli
31	702	29.4	595	14 US-10-052-092-13	Sequence 13, Appli
32	702	29.4	595	14 US-10-052-092-14	Sequence 14, Appli
33	702	29.4	595	14 US-10-207-655-61	Sequence 61, Appli
34	702	29.4	595	14 US-10-177-293-128	Sequence 128, App
35	702	29.4	595	14 US-10-157-899A-55	Sequence 55, Appli
36	702	29.4	595	14 US-10-437-107-9	Sequence 9, Appli
37	702	29.4	595	14 US-10-437-107-13	Sequence 13, Appli
38	702	29.4	595	14 US-10-437-107-14	Sequence 14, Appli
39	702	29.4	595	14 US-10-095-373A-2	Sequence 2, Appli
40	702	29.4	595	14 US-10-392-274-2	Sequence 2, Appli
41	702	29.4	595	15 US-10-148-835-6	Sequence 6, Appli
42	702	29.4	595	15 US-10-144-198-42	Sequence 42, Appli
43	701	29.4	595	15 US-10-148-835-2	Sequence 2, Appli
44	701	29.4	595	15 US-10-148-835-7	Sequence 7, Appli
45	697	29.2	595	15 US-10-148-835-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-10-054-841-4  
; Sequence 4, Application US/10054841  
; Publication No. US20030119100A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Fang  
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
; FILE REFERENCE: 19999YCA  
; CURRENT APPLICATION NUMBER: US/10/054, 841  
; CURRENT FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: 09/487, 379  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: 09/141, 000  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/078, 633  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/062, 902  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/057, 090  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Human  
; US-10-054-841-4

Query Match 100.0%; Score 2388; DB 14; Length 458;  
Best Local Similarity 100.0%; Pred. No. 2.4e-215;  
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEHELICRMSNKRHRIDSSCSFFIKTEPSSPASTLDSVNHSPGG 60  
Db 1 MDSVELCLPESFSLHYEHELICRMSNKRHRIDSSCSFFIKTEPSSPASTLDSVNHSPGG 60  
QY 61 SSDASGSYSTWNGHQNGLDSPPLYPAPILGSGPYRKLTYDDCSSTIVEDPQTKCEYML 120

Db 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120

QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFPKRTIQGNIEYSCPATNECEITKRRKSC 180

Db 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFPKRTIQGNIEYSCPATNECEITKRRKSC 180

QY 181 QACRFMKCLKVGMKEGVRLDRVGRGROKYKRRIDAENSPYLNPOLVOPAKKPYNKIVSH 240

Db 181 QACRFMKCLKVGMKEGVRLDRVGRGROKYKRRIDAENSPYLNPOLVOPAKKPYNKIVSH 240

QY 241 LIVAPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300

Db 241 LIVAPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300

QY 301 LOSAMWEILILGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDLNNAILOLVKKYKSMK 360

Db 301 LOSAMWEILILGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDLNNAILOLVKKYKSMK 360

QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVOKLODVLHEALODYEAGQHMEDPRRAGKMLMT 420

Db 361 LEKEEFVTLKAIALANSDSMHIEDVEAVOKLODVLHEALODYEAGQHMEDPRRAGKMLMT 420

QY 421 LPLLRQSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

Db 421 LPLLRQSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

RESULT 2

US-10-355-218-1

; Sequence 1, Application US/10355218

; Publication No. US20040009558A1

; GENERAL INFORMATION:

; APPLICANT: MORAS, DINO

; APPLICANT: RENAUD, JEAN-PAUL

; APPLICANT: GRESCHIK, HOLGER

; APPLICANT: WURTZ, JEAN-MARIE

; TITLE OF INVENTION: THE LIGAND BINDING POCKET PEPTIDE FRAGMENT OF THE

; TITLE OF INVENTION: ESTROGEN-RELATED RECEPTOR 3 (ERR3) AND USES THEREOF

; FILE REFERENCE: 0510-1060

; CURRENT APPLICATION NUMBER: US/10/355,218

; CURRENT FILING DATE: 2003-01-31

; PRIOR APPLICATION NUMBER: 60/352,551

; PRIOR FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 458

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-355-218-1

Query Match 100.0%; Score 2388; DB 15; Length 458;

Best Local Similarity 100.0%; Pred. No. 2.4e-215;

Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEBELLCRMSNKNDRHIDSSCSFRTKEPSSPASLTDVNHHSPPG 60

Db 1 MDSVELCLPESFSLHYEBELLCRMSNKNDRHIDSSCSFRTKEPSSPASLTDVNHHSPPG 60

QY 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120

Db 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120

QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFPKRTIQGNIEYSCPATNECEITKRRKSC 180

Db 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFPKRTIQGNIEYSCPATNECEITKRRKSC 180

QY 181 QACRFMKCLKVGMKEGVRLDRVGRGROKYKRRIDAENSPYLNPOLVOPAKKPYNKIVSH 240

Db 181 QACRFMKCLKVGMKEGVRLDRVGRGROKYKRRIDAENSPYLNPOLVOPAKKPYNKIVSH 240

QY 241 LIVAPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300

Db 241 LIVAPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300

QY 301 LOSAMWEILILGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDLNNAILOLVKKYKSMK 360

Db 301 LOSAMWEILILGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDLNNAILOLVKKYKSMK 360

QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVOKLODVLHEALODYEAGQHMEDPRRAGKMLMT 420

Db 361 LEKEEFVTLKAIALANSDSMHIEDVEAVOKLODVLHEALODYEAGQHMEDPRRAGKMLMT 420

QY 421 LPLLRQSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

Db 421 LPLLRQSTKAVQHFYNIKLEGVPMHKLFLMLEAKV 458

RESULT 3

US-10-054-841-6

; Sequence 6, Application US/10054841

; Publication No. US20030119100A1

; GENERAL INFORMATION:

; APPLICANT: Chen, Fang

; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR

; FILE REFERENCE: 19999YCA

; CURRENT APPLICATION NUMBER: US/10/054,841

; CURRENT FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: 09/487,379

; PRIOR FILING DATE: 2000-01-18

; PRIOR APPLICATION NUMBER: 09/141,000

; PRIOR FILING DATE: 1998-08-26

; PRIOR APPLICATION NUMBER: 60/078,633

; PRIOR FILING DATE: 1998-03-19

; PRIOR APPLICATION NUMBER: 60/062,902

; PRIOR FILING DATE: 1997-10-21

; PRIOR APPLICATION NUMBER: 60/057,090

; PRIOR FILING DATE: 1997-08-27

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 418

; TYPE: PRT

; ORGANISM: Human

US-10-054-841-6

Query Match 89.6%; Score 2139; DB 14; Length 418;

Best Local Similarity 100.0%; Pred. No. 5.4e-192;

Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEBELLCRMSNKNDRHIDSSCSFRTKEPSSPASLTDVNHHSPPG 60

Db 1 MDSVELCLPESFSLHYEBELLCRMSNKNDRHIDSSCSFRTKEPSSPASLTDVNHHSPPG 60

QY 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120

Db 61 SSDASGSYSTMNGHONGLDSPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPQTKCEYML 120

QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFPKRTIQGNIEYSCPATNECEITKRRKSC 180

Db 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFPKRTIQGNIEYSCPATNECEITKRRKSC 180

QY 181 QACRFMKCLKVGMKEGVRLDRVGRGROKYKRRIDAENSPYLNPOLVOPAKKPYNKIVSH 240

Db 181 QACRFMKCLKVGMKEGVRLDRVGRGROKYKRRIDAENSPYLNPOLVOPAKKPYNKIVSH 240

QY 241 LIVAPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300

Db 241 LIVAPEKIYAMPDPTVPDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTLSLADQMSL 300

QY 301 LOSAMWEILILGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDLNNAILOLVKKYKSMK 360

Db 301 LOSAMWEILILGVVYRSLSFEDELVYADDYIMDEDQSKLAGLLDLNNAILOLVKKYKSMK 360

QY 361 LEKEEFTLKAIALANSDSMHIEDVEAVQKLQDVLEHALQDYEAGQHME 409  
Db 361 LEKEEFTLKAIALANSDSMHIEDVEAVQKLQDVLEHALQDYEAGQHME 409

RESULT 4

US-10-054-841-2  
; Sequence 2, Application US/10054841  
; Publication No. US20030119100A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Fang  
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
; TITLE OF INVENTION: RECEPTOR PROTEINS  
; FILE REFERENCE: 1999YCA  
; CURRENT APPLICATION NUMBER: US/10/054,841  
; CURRENT FILING DATE: 2002-01-23  
; PRIOR APPLICATION NUMBER: 09/487,379  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: 09/141,000  
; PRIOR FILING DATE: 1998-08-26  
; PRIOR APPLICATION NUMBER: 60/078,633  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: 60/062,902  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/057,090  
; PRIOR FILING DATE: 1997-08-27  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Human  
US-10-054-841-2

Query Match 75.3%; Score 1799; DB 14; Length 500;  
Best Local Similarity 78.1%; Pred. No. 6.5e-160;  
Matches 339; Conservative 50; Mismatches 43; Indels 2; Gaps 2;

QY 24 MSNKDRHIDSSCGSFITKEPSSPASLTDSVNHSPGGSSDASGSYSTMNGHQGLDSP 83  
Db 1 MSSDDRHLGSSCGSFITKEPSSPGSIDALSHSPSSGSDASGGFGALGTHANGLDSP 60  
QY 84 LYPAPILGGSGPVKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143  
Db 61 MFAGAG-LGGT-PCRKSIEDCASGIMEDSAIKCEYMLNAI PKRLCLVCGDIASGYHYGVA 118  
QY 144 SCEACKAFFKRTIQGNI EYSCPATNECEITKRRRKSQAACRFMKCLKVGMKEGVRLDRV 203  
Db 119 SCEACKAFFKRTIQGNI EYSCPATNECEITKRRRKSQAACRFMKCLKVGMKEGVRLDRV 178  
QY 204 RGRQKYKRRIDAENSPYLNLPOLVQPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK 263  
Db 179 RGRQKYKRRILDSESSPYLSLQISPPAKKPLTKIVSYLLVAEPDKLYAMPPEGMPEGDIK 238  
QY 264 ALTTLCDLADRELVIIGWAKHIPGFSTLSLADQMSLLQSAWMEILLGVYRSLSEFEDE 323  
Db 239 ALTTLCDLADRELVIIGWAKHIPGFSTLSLADQMSLLQSAWMEILLGVYRSLSEFEDE 298  
QY 324 LVYADDYIMDEDOSKLAGLLDNNAILQLVKKYKSMKLEKEEFVTLKAIALANSDSMHIE 383  
Db 299 LVYAEADYIMDEHRSRLAGLLLELYRAILLQVRRYKKLKVEKEEFVTLKAIALANSDSMYIE 358  
QY 384 DVEAVQKLQDVLEHALQDYEAGQHMEPPRACKMLMTLLPLRQSTKAVQHFYNIKLEGG 443  
Db 359 DLEAVQKLQDLHEALQDYELSQRHEEPWRTGKLLLTPLLRQTAKAQVQHFYSVKLQGG 418  
QY 444 VPMHKLFLMLEAK 457  
Db 419 VPMHKLFLMLEAK 432

RESULT 5  
US-10-278-481-19

; Sequence 19, Application US/10278481  
; Publication No. US20030113803A1  
; GENERAL INFORMATION:  
; APPLICANT: KARO BIO AB  
; TITLE OF INVENTION: Orphan receptor  
; NUMBER OF SEQUENCES: 19  
; STREET: c/o Center for Biotechnology and Department  
; of Medical Nutrition, Karolinska Nutrition, Karolinska  
; Institute  
; CITY: Huddinge  
; COUNTRY: Sweden  
; ZIP: S-14186  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/278,481  
; FILING DATE: 23-Oct-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/333,057  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/836,620  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: GB 9518272.1  
; FILING DATE: 08-SEP-1995  
; APPLICATION NUMBER: GB 9605550.4  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: GB 9607532.0  
; FILING DATE: 11-APR-1996  
; APPLICATION NUMBER: GB 9609576.5  
; FILING DATE: 08-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GUSTAFSSON, Jan-Ake  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 431 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-10-278-481-19

Query Match 74.3%; Score 1775; DB 14; Length 431;  
Best Local Similarity 76.4%; Pred. No. 9.3e-158;  
Matches 331; Conservative 51; Mismatches 49; Indels 2; Gaps 1;

QY 24 MSNKDRHIDSSCGSFITKEPSSPASLTDSVNHSPGGSSDASGSYSTMNGHQGLDSP 83  
Db 1 MSSEDRHLGSSCGSFITKEPSSPGSIDALSHSPSSGSDASGGFGMALGTHANGLDSP 60  
QY 84 LYPAPILGGSGPVKLYDDCSSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143  
Db 61 MFAGAGL--GGNPCRKSIEDCTSGIMEDSAIKCEYMLNAI PKRLCLVCGDIASGYHYGVA 118  
QY 144 SCEACKAFFKRTIQGNI EYSCPATNECEITKRRRKSQAACRFMKCLKVGMKEGVRLDRV 203  
Db 119 SCEACKAFFKRTIQGNI EYSCPATNECEITKRRRKSQAACRFMKCLKVGMKEGVRLDRV 178  
QY 204 RGRQKYKRRIDAENSPYLNLPOLVQPAKKPYNKIVSHLLVAEPEKIYAMPDPTVPDSDIK 263  
Db 179 RGRQKYKRRILDSENSPYLSLQISPPAKKPLTKIVSYLLVAEPDKLYAMPDDVPEGDIK 238  
QY 264 ALTTLCDLADRELVIIGWAKHIPGFSTLSLADQMSLLQSAWMEILLGVYRSLSEFEDE 323  
Db 239 ALTTLCDLADRELVIISWAKHIPGFSNLTLDQMSLLQSAWMEILLGVYRSLPYDDK 298  
QY 324 LVYADDYIMDEDOSKLAGLLDNNAILQLVKKYKSMKLEKEEFVTLKAIALANSDSMHIE 383  
Db 299 LAYAEADYIMDEHRSRLVGLLELYRAILLQVRRYKKLKVEKEEFVTLKAIALANSDSMYIE 358



```
QY      384 DVEAVÖKLÖDVILHEALÖDYEAGQMEDPRRAQKMLMTLPBLRÖTSTKAVQHFINIKLEGG 443  
        ::|||||::|||::|::|::|::|::|::|::|::|::|::|::|::|  
Db      359 NIEAVÖKLÖDVLHEALÖDYELSGRHEEBRRARGLTLTPPLRQTAKAAVQHFFYSVKLQGK 418  
  
QY      444 VPMHKLFLJEMLEA 456  
        |||||||  
Db      419 VPMHKLFLJEMLEA 431
```

## RESULT 6

```

US-10-297-653-1
; Sequence 1, Application US/10297653
; Publication No. US20030215829A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: CHINN, Anna M.
; APPLICANT: LU, Yan
; APPLICANT: TRIBOULEY, Catherine M.
; TITLE OF INVENTION: NUCLEAR HORMONE RECEPTORS
; FILE REFERENCE: PI-0119 USN
; CURRENT APPLICATION NUMBER: US/10/297,653
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/209,863
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030215829A1 7475568CD1
US-10-297-653-1

```

Query Match	52.7%;	Score 1259;	DB 15;	Length 486;
Best Local Similarity	54.9%;	Pred. No. 3.6e-109;		
Matches 256; Conservative	67;	Mismatches 79;	Indels 64;	Gaps 8;

[illegible]

## RESULT 7

```

US-10-211-239-1
; Sequence 1, Application US/10211239
; Publication No. US20030103965A1
; GENERAL INFORMATION:
; APPLICANT: Jung, Birgit
; APPLICANT: Kraut, No. US20030103965Albert
; APPLICANT: Mueller, Stefan
; TITLE OF INVENTION: Method for Identifying Substances Which Positively
; TITLE OF INVENTION: Influence Inflammatory Conditions
; FILE REFERENCE: 0652.2340001
; CURRENT APPLICATION NUMBER: US/10/211,239
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/315,775
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: EP 01119003.0
; PRIOR FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-211-239-1

```

Query Match	52.7%;	Score 1258.5;	DB 14;	Length 519;
Best Local Similarity	55.1%;	Pred. No. 4.4e-109;		
Matches 256;	Conservative 67;	Mismatches 79;	Indels 63;	Gaps 8;

[illegible]

## RESULT 8

US-10-278-481-18  
; Sequence 18, Application US/10278481  
; Publication No. US20030113803A1  
; GENERAL INFORMATION:  
; APPLICANT: KARO BIO AB  
; TITLE OF INVENTION: Orphan receptor  
; NUMBER OF SEQUENCES: 19  
; STREET: c/o Center for Biotechnology and Department  
; of Medical Nutrition, Karolinska Nutrition, Karolinska  
; Institute  
; CITY: Huddinge

```

; COUNTRY: Sweden
; ZIP: S-14186
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/278,481
; FILING DATE: 23-Oct-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,057
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/836,620
; FILING DATE: <Unknown>
; APPLICATION NUMBER: GB 9518272.1
; FILING DATE: 08-SEP-1995
; APPLICATION NUMBER: GB 9605550.4
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: GB 9607532.0
; FILING DATE: 11-APR-1996
; APPLICATION NUMBER: GB 9609576.5
; FILING DATE: 08-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: GUSTAFSSON, Jan-Ake
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-278-481-18

Query Match      50.8%; Score 1212.5; DB 14; Length 518;
Best Local Similarity 53.5%; Pred. No. 9.3e-105;
Matches 250; Conservative 67; Mismatches 85; Indels 65; Gaps 9;

QY 22 CRMSNKRHHIDSSCSS-----FIKTEPSSPASLTDSVNHSPGSSSD-----A 64
Db 85 CPTALPEPQVTSAMSSQVVGIEPLVTKAEPASP-----DSPKSSSETEPVALA 135
QY 65 SG-SYSTMNGHQNGLDSPPLPSAPILIGSGSPVRKLYDDCSSTIVEDPQTKCEYMNISM 123
Db 136 PGPAPTKCLPGHKEBEDGE-----GAGP-----GEQGGKLVLSL 171
QY 124 PKRLCTVCGDIASGYHYGVASCACAKAFKKRTIOGNIYESCPATNECEITKRRKSCQAC 183
Db 172 PKRLCTVCGDVASGYHYGVASCACAKAFKKRTIOGSIYESCPASNECEITKRRKACQAC 231
QY 184 REMKCLVGMKEGVRLDRVGRGQKKRRIDAENSPYLNQVQP-----AKRPY 234
Db 232 RFTKCTRVGMKEGVRLDRVGRGQKKRRPEVDPLFPGPFPAGPLAVAGGPRKTAAPV 291
QY 235 NKIVSHLLVAEPEKIYAMPDPVDSDIKALTTLCDLADRELVTIIGWAKHIPGFSTISL 294
Db 292 NALVSHLLIVEPEKLYAMPDPAGDGHLPVAVATLCDLFDREIVVTISWAKSIPGFSSLSL 351
QY 295 ADQMSLLQSAWMEILLGVVYRSLSFEDELVYADDYIMDEDSKLAGLLDNNAILQIVK 354
Db 352 SDQMSVLQSVWMEVLVLGVAQRSLPLQDELAFADLVLTIEGARAAAGLGELGAALLQIVR 411
QY 355 KYKSMKLEKEEFVTIKALALANSDSMHIEDVEAV-QKLQDVLHEALQDYEAGQ---HME 409
Db 412 RLQALRLERBEVYLKALALANSDSVHIEDEPRLWSCEKLLHEALLLEYEAGRAGPGGA 471
QY 410 DPRRAGKMLMTPLRLQSTKAQVQHFYNIKLEGVPMHKLFLMLEA 456
Db 472 FRRRAGRLLLTPLRLQSTKAQVLAHFYGVKLEGVPMHKLFLMLEA 518
```

RESULT 9

```

US-10-093-463-202
; Sequence 202, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glenda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Paturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypeptic
; TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; PRIOR APPLICATION NUMBER: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
```

PRIOR APPLICATION NUMBER: 60/275,578  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 370  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 202  
LENGTH: 516  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-093-463-202

Query Match 44.9%; Score 1073; DB 15; Length 516;  
Best Local Similarity 50.4%; Pred. No. 1.2e-91;  
Matches 226; Conservative 68; Mismatches 98; Indels 56; Gaps 9;

QY 31 IDSGCSFIKTEPSSPASLTDSVNHSPGSSD-----ASG-SYSTMNGHONGLD 80  
DB 100 VGSEPLQYIMAEPAAP-----DSPKSSSETETEPVAPALAPGAPATHCLPGHKEED 150  
QY 81 SPPLPSAPILGSGSPVRKLYDDCSSTIVEDPQTCEYMLNSMPKRLCLVCGDIAGYHY 140  
DB 151 GE-----GAGP-----GEOGGKLVLSLSKRLCLVCGDVASGYHC 186  
QY 141 GVASCEACAKAFKRTIQGNIETSCPAETNECEITKRRKSCQACREMKCLKVGMKEGVRL 200  
DB 187 GVSSCEDCKAFKRTIQGSMETSCLASNECEITKRRKACQACRFTKSLRLHQEPAGARL 246  
QY 201 DRVGGGQKRRIDAENSPY-----LNPOLVQPAK--PYNKIVSHLLVAEPEKIYAM 252  
DB 247 DRVGGGQKRRIDAENSPY-----LNPOLVQPAK--PYNKIVSHLLVAEPEKIYAM 252  
QY 253 PDPVTPSDIKALITLCLDLADRELVIITGAKHIFGFTSLSLADQMSLIQSAMMEILIG 312  
DB 306 PDPAGPDGHLPAVATLCLDLFDEIIVTISWAKSIPFSSLSLSDQMSVILQVWMEVLVPG 365  
QY 313 VVYRSLSFEDLVYADDYIMDEDQSKLAGLIDNNAILQLVKYKSMKEKEEFVTLKAI 372  
DB 366 VAQRSLPLQDELAFAEDLVLDEEGARAAGLGELGALLQLVRLQSLRERGEYVLLKAL 425  
QY 373 ALANSDSMHIEDVEAVOKLQDVLHEALQDYEAQO---HMEDPRRAGKMLTLPILRQTS 428  
DB 426 ALANSDSVPIEDAEVQLPEAPHEALLLEYEAGRAAGTGGAEERRRPRLFTLLPLHQTA 485  
QY 429 TKAVQHPYNIKLEGVPMHKLFLMLEA 456  
DB 486 GKVLAFHYGVKLEGVPMHKLFLMLEA 513

RESULT 10  
US-10-355-218-8  
Sequence 8, Application US/10355218  
Publication No. US20040009558A1  
GENERAL INFORMATION:  
APPLICANT: MORAS, DINO  
APPLICANT: RENAUD, JEAN-PAUL  
APPLICANT: GRESCHIK, HOLGER  
APPLICANT: WURTZ, JEAN-MARIE  
TITLE OF INVENTION: THE LIGAND BINDING POCKET PEPTIDE FRAGMENT OF THE  
FILE REFERENCE: 0510-1060  
CURRENT APPLICATION NUMBER: US/10/355,218  
CURRENT FILING DATE: 2003-01-31  
PRIOR APPLICATION NUMBER: 60/352,551  
PRIOR FILING DATE: 2002-01-31  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 230  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-355-218-8

Query Match 39.7%; Score 947; DB 15; Length 230;  
Best Local Similarity 79.4%; Pred. No. 2.4e-80;

Matches 181; Conservative 29; Mismatches 18; Indels 0; Gaps 0;  
QY 230 AKPPYNKIVSHLLVAEPEKIYAMPDPTVPDSIDIKALITLCLDLADRELVIITGAKHIFG 289  
DB 1 AKKPLTKIVSYLLVAEPEKLYAMPPEGMEGDIKALITLCLDLADRELVIITGAKHIFG 60  
QY 290 STLADQMSLIQSAMMEILIGVYRSLSFEDLVYADDYIMDEDQSKLAGLIDNNAIL 349  
DB 61 SLSLGDQMSLIQSAMMEILIGVYRSLPYDDKLVYADDYIMDEHSRLAGLLELYRAI 120  
QY 350 LQIVKYSKMKLEKEEFVTLKALALANSDSMHIEDVEAVOKLQDVLHEALQDYEAQOME 409  
DB 121 LQIVRYKKLKEKEEFVTLKALALANSDSMYIEDLEAVOKLQDVLHEALQDYELSQRHE 180  
QY 410 DRRRAGKMLTLPILRQTSKAVQHPYNIKLEGVPMHKLFLMLEAK 457  
DB 181 EPWRTGKLLITLPLLRQTAQAVQHFYSVKLQGVPMHKLFLMLEAK 228

RESULT 11  
US-10-355-218-9  
Sequence 9, Application US/10355218  
Publication No. US20040009558A1  
GENERAL INFORMATION:  
APPLICANT: MORAS, DINO  
APPLICANT: RENAUD, JEAN-PAUL  
APPLICANT: GRESCHIK, HOLGER  
APPLICANT: WURTZ, JEAN-MARIE  
TITLE OF INVENTION: THE LIGAND BINDING POCKET PEPTIDE FRAGMENT OF THE  
FILE REFERENCE: 0510-1060  
CURRENT APPLICATION NUMBER: US/10/355,218  
CURRENT FILING DATE: 2003-01-31  
PRIOR APPLICATION NUMBER: 60/352,551  
PRIOR FILING DATE: 2002-01-31  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 234  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-355-218-9

Query Match 31.6%; Score 755; DB 15; Length 234;  
Best Local Similarity 62.7%; Pred. No. 2.8e-62;  
Matches 143; Conservative 43; Mismatches 38; Indels 4; Gaps 1;

QY 233 PYNKIVSHLLVAEPEKIYAMPDPTVPDSIDIKALITLCLDLADRELVIITGAKHIFGFTSL 292  
DB 4 PYNALVSHLLVAEPEKLYAMPDPAGPDGHLPAVATLCLDLFDEIIVTISWAKSIPGFSSSL 63  
QY 293 SLADQMSLIQSAMMEILIGVYRSLSFEDLVYADDYIMDEDQSKLAGLIDNNAILQL 352  
DB 64 SLSDQMSVLQSVWMEVLVGVAAQRSLPLQDELAFAEDLVLDEEGARAAGLGELGALLQL 123  
QY 353 VKKYSKMKLEKEEFVTLKALALANSDSMHIEDVEAVOKLQDVLHEALQDYEAQO---HM 408  
DB 124 VRRLOALRLEREYVLLKALALANSDSVHIEDAEVQLREALHALLLEYEAGRAGGCG 183  
QY 409 EDPRRAGKMLTLPILRQTSKAVQHPYNIKLEGVPMHKLFLMLEA 456  
DB 184 AERRRAGRLITLPLLRQTAGKVLAFHYGVKLEGVPMHKLFLMLEA 231

RESULT 12  
US-10-450-484A-1  
Sequence 1, Application US/10450484A  
Publication No. US20040121355A1  
GENERAL INFORMATION:  
APPLICANT: Sumitomo Chemical Company Limited  
TITLE OF INVENTION: Estrogen Receptor Genes and Utilization Thereof  
FILE REFERENCE: 600630-5US (560178)  
CURRENT APPLICATION NUMBER: US/10/450,484A



;; CURRENT FILING DATE: 2004-02-09  
;; PRIOR APPLICATION NUMBER: JP 2000/379993  
;; PRIOR FILING DATE: 2000-12-14  
;; PRIOR APPLICATION NUMBER: PCT/JP01/09996  
;; PRIOR FILING DATE: 2001-11-15  
;; NUMBER OF SEQ ID NOS: 26  
;; SOFTWARE: Patentin version 3.2  
;; SEQ ID NO 1  
;; LENGTH: 587  
;; TYPE: PRT  
;; ORGANISM: Caiman crocodilus  
;; US-10-450-484A-1

Query Match 30.4%; Score 726; DB 16; Length 587;  
Best Local Similarity 38.2%; Pred. No. 6.1e-59;  
Matches 175; Conservative 80; Mismatches 143; Indels 60; Gaps 14;

QY 38 FIKTEPSSPASLTDSVNHSPGG---SSDASGSYSTMNGHONGDSPPLPSAPILGG 93  
DB 104 FLQTAP-----QLSPFVHHSSQVPPYLENDQSGF-----GMRASSTFYRPSADSRLHQ 153  
QY 94 SGFVRKLYDDCSSSTIVEDPQTKCEYMLNSMPK-RLCLVCGDIASGYHYGVAACEACKAF 152  
DB 154 SGRER-----MSST-----SEKASLSESTKTRCYCAVCNDYASGYHYGVAWSCGCKAF 203  
QY 153 KRTIQNIIEYSCPATNECEITKRRKSCQACRFMKCLKVGMLEKGVRLDRVGRQ-KYK 211  
DB 204 KRSTQGHNDYMCPATNQCTIDKNRRKSCQACRLKCYEYGMKGIRKDR-RGGRMLKQK 262  
QY 212 RR---IDAENSP-----YLNQVLQPAKK-----PYNKIVSHLLVAEPEKIY 250  
DB 263 RQREODARNGETATAEMRPTPLWTSPLYIKHTKNSPALSLTAEQWVSALLAEPPIVY 322  
QY 251 AMPDPTEVDSDIKALTTLCDLADRELVIYIGWAKHIPGFSTLSLADQMSLQSAWMEILI 310  
DB 323 SEYDNRPFNEASMTLLTNLADRELVIYMINWAKRVPGEVDLTLDQVHLLECAWLEILM 382  
QY 311 LGVYRSLSEFEDLVYADYIMDEDQSK-LAGLLDLNNAILOVKKYKSMLEKEEFTVL 369  
DB 383 IGLVWRSMHEPGKLLFAPNLLDRNQKCEGMEVEIFDMLLATARFRMNLQGEETVCL 442  
QY 370 KATALAN-----SDSMHIEDVEAVQKLODVLEALQDYEAQOM--EDPRRAGKML 418  
DB 443 KSTLLNSGVYTFPLSTKLSLEKDYIHRVLDKITDTLHLMKXSGLSLQOQHRRLAQLL 502  
QY 419 MTLPLRQSTKAVQHFYNIKLEKGVPMKLFLEMLEA 456  
DB 503 LILSHIRMSNKGMEHLYNMCKKNVVPYDILLLEMLDA 540

RESULT 13

US-10-278-481-16  
; Sequence 16, Application US/10278481  
; Publication No. US20030113803A1  
; GENERAL INFORMATION:  
; APPLICANT: KARO BIO AB  
; TITLE OF INVENTION: Orphan receptor  
; NUMBER OF SEQUENCES: 19  
; STREET: c/o Center for Biotechnology and Department  
; of Medical Nutrition, Karolinska Nutrition, Karolinska  
; Institute  
; CITY: Huddinge  
; COUNTRY: Sweden  
; ZIP: S-14186  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/278,481  
; FILING DATE: 23-Oct-2002  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/333,057  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/836,620  
FILING DATE: <Unknown>  
APPLICATION NUMBER: GB 9518272.1  
FILING DATE: 08-SEP-1995  
APPLICATION NUMBER: GB 9605550.4  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: GB 9607532.0  
FILING DATE: 11-APR-1996  
APPLICATION NUMBER: GB 9609576.5  
FILING DATE: 08-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: GUSTAFSSON, Jan-Ake  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 596 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
ORIGINAL SOURCE:  
ORGANISM: Rattus rattus  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-10-278-481-16

Query Match 30.0%; Score 716.5; DB 14; Length 596;  
Best Local Similarity 39.2%; Pred. No. 4.9e-58;  
Matches 168; Conservative 82; Mismatches 128; Indels 51; Gaps 13;

QY 73 NGHONGDSPPLY---PSAPILGSGSP---VRKLYDDCSSSTIVEDPQTKCE---YMLN 121  
DB 129 HGHQ-----VPYILENPESAVAVRDTGPPAFYRSNSDNRRQNGERLSSSEKGMIMES 183  
QY 122 SMPKRLCLVCGDIASGYHYGVAACEACKAFKRTIQNIIEYSCPATNECEITKRRKSCQ 181  
DB 184 AKETRYCAVCNDYASGYHYGVAWSCGCKAFKRSIQGHNDYMCPATNQCTIDKNRRKSCQ 243  
QY 182 ACFEMCLKVGMLEKGVRLDRVGRQ-KYK-RIDAEN-----SPYLNQVL 226  
DB 244 ACRLRKCYEYGMKGIRKDR-RGGRMLKHKRQRDDLEGRNEMGTSGDMRAANLWPSPLV 302  
QY 227 VQPAKK-----PYNKIVSHLLVAEPEKIYAMPDPTEVDSDIKALTTLCDLADRELVI 279  
DB 303 IKHTKNSPALSLTADQWVSALLDAEPPLIYSEYDPSRPFSEASMGLLTNLADRELVIH 362  
QY 280 IGWAKHIPGFSTLSLADQMSLQSAWMEILIGVYRSLSEFEDLVYADYIMDEDQSK- 338  
DB 363 INWAKRVPGEFDLTHDQVHLLECAWLEILMIGLVWRSMHEPGKLLFAPNLLDRNQKC 422  
QY 339 IAGLLDLNNAILOVKKYKSMLEKEEFTVLKATALAN-----SDSMHIEDVEAVQK 390  
DB 423 VEGMEVEIFDMLLATSSRFMYNLQGEETVCLKSTLLNSGVYTFPLSTKLSLEKDHHR 482  
QY 391 LODVLHEALQDY--EAGGHMEDP-RRAGKMLMTPLRQSTKAVQHFYNIKLEKGVPMH 447  
DB 483 VLDKINDTLHLMAKAGLTIQOQHRRLAQLLILSHIRMSNKGMEHLYNMCKKNVVPY 542  
QY 448 KLFLEMLEA 456  
DB 543 DILLEMLDA 551

RESULT 14

US-10-451-768A-1  
; Sequence 1, Application US/10451768A  
; Publication No. US20040115765A1  
; GENERAL INFORMATION:  
; APPLICANT: Sumitomo Chemical Company Limited  
; TITLE OF INVENTION: Estrogen receptor gene and its use  
; FILE REFERENCE: 560163  
; CURRENT APPLICATION NUMBER: US/10/451,768A  
; PRIOR FILING DATE: 2003-12-24  
; PRIOR APPLICATION NUMBER: JP 2000/392262  
; PRIOR FILING DATE: 2000-12-25

```

; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 1
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Crnemidophorus uniparens
;
US-10-451-768A-1

```

Query Match	29.9%;	Score 715;	DB 16;	Length 581;
Best Local Similarity	34.9%;	Pred. No. 6.5e-58;		
Matches 164;	Conservative 80;	Mismatches 133;	Indels 93;	Gaps 10;

```

QY 30 HIDSCESSFKTEPPSPASLT-----DSVNHSPPGSSSDASGSYSSTMNGHONGL 79
      | : :: ||||| |
Db 117 HHNQVPPYYIENEPSSSAMREAFPTAFYRPGSENRRHHGGRASNSEKGSLS----- 166

QY 80 DSPPLYPSPAPIIGSGSPVRKLYDDCSSITVEDPQTKECYMLNSMPKRLCLVCGDIASGYH 139
      | : :: ||||| |
Db 167 -----MESTKETRYCAVCNDYASGYH 187

QY 140 YGVASCEACKAFKKRTIOGNIEYSCPATNCEITKRRKSCQACREMKLVGMKEGVR 199
      || || || || || || || : : || || || || || || || || || || : : ||
Db 188 YGVWSCGCKAFKKRSIOGHNDYMCPATNQCTIDKNRRKSCQACRLRKCYEVGMMKGIR 247

QY 200 LDRVRGGRQ-KYKRRID-----AENSP-YLNPQLVQPAKK-----PYNKIV 238
      || || || | : || : | : : : : : : : : : : : : : : : : : : : : :
Db 248 KDR-RGGRMLKHKRQDELDGRNAVAVTEARNTTLPSPLMIKHKSXNSPALSLTAEQMV 306

QY 239 SHLLVAEPEKIYAMPDPTVPBDSIDIKALTTCDLADRELVIIGWAKHIFGSTLSLADQM 298
      || || || | : : : : : : : : : : : : || || || || || || : : ||
Db 307 SALLDAEPPIVSEYDPSPSPSEASVMTLLTNLADRELVHMITWAKRVPGEVDLALHDQV 366

QY 239 SLLQSAWMEILLGVVYRSLSFEDELVYADDTIMDEDQSK-LAGLLDLNNAILLQLVKKYK 357
      || : : || : || : || : || : : || : : || : : || : : || : : || : :
Db 367 HILBCAWLEILMIGLWRSLEHPGKLFPANLLDRSQGMCEGFEVEIDMILLATSSRFR 426

QY 358 SMKLEKEEFVTLKAIALAN-----SDSMHTEDEVAOVOKLQDVLHEALQDYEAQOHM- 408
      | : : || || || || || | | : : || : : || : : || : : || : : || : :
Db 427 MNNTIQEEFVCKLSIILNSGIYTFLSSTLRSLERKEHIIHRVLDKITDTITLHMAKSGLS 486

QY 409 --EDPRRAGKMTMLPLLRQTSKAVQHFYNIKLEGKVPMAHKLFLMLEA 456
      : : || : : || : : || : : || : : || : : || : : || : : || : : ||
Db 487 LQQQHRRLAQLLMLSHIRHMSNKGMEHLYNMKCKNAVVPDYDLLEMLDA 536

```

[illegible]

```
Search completed: August 3, 2004, 12:43:43
Job time : 46 secs
```

Query Match	29.9%;	Score 713.5;	DB 14;	Length 595;
Best Local Similarity	42.3%;	Pred. No. 9.4e-58;		
Matches 159; Conservative	69;	Mismatches 105;	Indels 43;	Gaps 10;